



UNIVERSIDAD CARLOS III DE MADRID

TESIS DOCTORAL

SISTEMA AVANZADO DE ASISTENCIA A LA CONDUCCION
MEDIANTE VISION POR COMPUTADOR PARA LA
DETECCION DE LA SOMNOLENCIA

Autor:

Marco Javier Flores Calero

Directores:

Dr. José María Armingol Moreno

Dr. Arturo de la Escalera Hueso

Departamento de Ingeniería de Sistemas y Automática

Leganés, Septiembre de 2009

TESIS DOCTORAL

SISTEMA AVANZADO DE ASISTENCIA A LA CONDUCCION MEDIANTE VISION POR COMPUTADOR PARA LA DETECCION DE LA SOMNOLENCIA

Autor: Marco Javier Flores Calero

Directores: Dr. José María Armingol Moreno

Dr. Arturo de la Escalera Hueso

Firma del Tribunal Calificador:

Firma

Presidente:

Vocal:

Vocal:

Vocal:

Secretario:

Calificación:

Leganés, Septiembre de 2009

Agradecimientos

En la era de la competencia y del progreso tecnológico, la prosperidad ya no se construye sobre reservas abundantes de mano de obra no cualificada, sino sobre la base de una fuerza de trabajo creativa, competente, que produce ideas, bienes y servicios intensivos en conocimiento, es así que en la actualidad, el capital intelectual es entendido como “la forma última de ventaja comparativa”.

En este sentido, la experiencia le ha enseñado al mundo que no hay desarrollo sin capital humano, y tal como afirma la Organización para la Cooperación y el Desarrollo Económico (OCDE), la creación, diseminación y aplicación del conocimiento seguirá siendo el principal motor del crecimiento de la economía mundial por muchos años más, por ello, es preciso extraer lo mejor de las potencialidades a través de la investigación.

El trabajo de investigación ha sido la base para el desarrollo de la sociedad contemporánea debido a que genera conocimientos frescos que permiten innovar, y a la vez, buscar *nuevas verdades* y desechar otras. Sin embargo, este motor de la innovación requiere el esfuerzo de muchos actores y el consumo de mucho tiempo y recursos, tanto económicos como personales, por lo que se constituye en una tarea ardua pero gratificante.

A lo largo de esta investigación varias personas han colaborado de forma anónima en su desarrollo, en especial, los miembros del *Departamento de Ingeniería de Sistemas y Automática* de la UC3M, a todos ellos deseo hacerles llegar mis más sinceros agradecimientos.

En especial, deseo agradecer a mis tutores Arturo y José María por sus acertados consejos y por dedicar parte de su valioso tiempo a discutir y analizar varios de los temas

de la tesis, así como por sus aportaciones y continuo espíritu de crítica constructiva. Pero de sobremanera, les agradezco por brindarme su confianza, apoyo y amistad durante el tiempo que ha durado esta investigación.

Finalmente, deseo agradecer de todo corazón a mi familia por su apoyo incondicional manifestado, a través, de sus palabras de ánimo que en todo este tiempo me han servido de sostén para seguir adelante con el trabajo, a pesar de la distancia física existente entre España y Ecuador.

Marco Javier Flores Calero

Leganés, Junio de 2009

Resumen

Esta tesis doctoral presenta un sistema para la detección de la somnolencia del conductor, basado en el análisis de los ojos. El sistema tiene la capacidad de adaptarse a cualquier persona, trabaja en tiempo real, bajo condiciones variables de iluminación y reales de conducción, generando en cada instante un índice de somnolencia que mide el estado de vigilia del conductor. Complementariamente, verifica la distracción analizando la orientación del rostro.

Los ojos generan una considerable cantidad de información del estado de una persona, son capaces de expresar alegría, tristeza, cansancio, etc., en particular, en el caso del conductor, indican su estado de atención y vigilia frente al volante, por ello, el principal objetivo de este trabajo ha sido determinar su posición, para continuar con su análisis hasta llegar a construir un índice de somnolencia.

En este caso, el índice implementado para detectar la somnolencia es el denominado PERCLOS; el mismo que ha sido validado, en estudios anteriores, como el más eficaz para estudiar esta problemática bajo la metodología aplicada en este trabajo. Para verificar la distracción se ha utilizado la información de la posición de los ojos y de la orientación del rostro.

Debido a la complejidad del problema para detectar y analizar los ojos, ha sido necesario dividirlo en dos partes: un sistema para trabajar durante el día y otro para trabajar durante la noche, en relación directa con la conducción diurna y nocturna; procurando de esta manera, cubrir el amplio espectro de ambientes donde se desarrolla la conducción.

Los sistemas planteados utilizan modernas técnicas de procesamiento de imágenes e Inteligencia Artificial, que los dotan de la suficiente capacidad para operar automática y eficazmente frente a las señales de somnolencia y distracción.

El sistema desarrollado para el día utiliza una cámara estándar que captura imágenes con iluminación natural, mientras que el sistema desarrollado para la noche utiliza un complejo sistema de visión basado en la iluminación infrarroja del tipo cercano, de esta manera, se aprovechan las propiedades espectrales para resaltar las pupilas de los ojos del conductor en condiciones de baja iluminación, como es el caso de la noche.

En varios experimentos realizados, sobre el vehículo experimental Ivvi, el sistema propuesto ha presentado magníficos resultados respecto a los objetivos planteados, y los inconvenientes han sido superados con éxito. Los experimentos han sido desarrollados en dos escenarios: (i) situaciones reales de conducción, y (ii) condiciones de laboratorio.

Finalmente, se puede resaltar que en este trabajo se ha presentado una nueva y completa propuesta que puede contribuir significativamente al desarrollo de la tecnología aplicada a prevenir los accidentes de tráfico, en el contexto, de los sistemas de seguridad activa.

Índice general

Agradecimientos	III
Resumen	V
Índice general	XII
Índice de figuras	XVIII
Índice de tablas	XIX
1. Introducción	1
1.1. Motivación	1
1.2. Somnolencia y distracción	3
1.2.1. Definición de somnolencia	3
1.2.2. Definición de distracción	4
1.2.3. Factores que producen somnolencia y sus consecuencias	4
1.2.4. Características visuales que indican la presencia de somnolencia . . .	4
1.3. La industria automotriz, su desarrollo, evolución y consecuencias	5
1.3.1. Producción mundial de vehículos	7
1.3.2. Congestión vehicular y contaminación del aire	8
1.4. Accidentes de tráfico	8
1.4.1. División de los accidentes de tráfico	10
1.4.2. Accidentes de tráfico por somnolencia	12

1.4.3.	Consejos para evitar la somnolencia durante la conducción	16
1.4.4.	Somnolencia y distracción como factores de accidentes de tráfico . .	16
1.5.	Objetivos	17
1.6.	Estructura de la tesis	18
2.	Detección de la somnolencia: Estado del Arte y Perspectivas	23
2.1.	Métodos utilizados para la detección de la somnolencia	23
2.1.1.	Análisis basados en patrones de conducción	24
2.1.2.	Análisis basados en los cambios físicos de los ojos y las expresiones faciales usando procesamiento de imágenes	25
2.1.3.	Análisis basados en el cambio de las medidas fisiológicas	26
2.2.	Métodos basados en el análisis de los ojos y expresiones faciales	27
2.2.1.	Métodos en el espectro visible	28
2.2.2.	Métodos en el infrarrojo cercano	32
2.3.	Sistemas comerciales	36
2.3.1.	Attention Technologies	36
2.3.2.	Seeing Machine	37
2.3.3.	Grupo Bosh	38
2.3.4.	Volvo Cars	38
2.3.5.	Mercedes-Benz Anti-Fatigue Driver Alarm	39
2.3.6.	AWAKE	40
2.3.7.	Mitsubishi Motors Corporation	41
2.3.8.	AssistWare Technology	42
2.3.9.	Delta Tooling	43
2.4.	Limitaciones de los sistemas comerciales actuales	44
2.5.	Conclusiones	44
3.	Apreciación Global y Sistema de Percepción	47
3.1.	Apreciación global y características generales	47
3.2.	Sistema de percepción del vehículo Ivvi	50
3.2.1.	Sistema de visión en el espectro visible	50

3.2.2. Sistema de visión en el infrarrojo cercano	50
3.3. Conclusiones	59
4. Detección de la Somnolencia en el Espectro Visible	61
4.1. Descripción del sistema en el espectro visible: diseño y alcance	61
4.2. Detección del rostro del conductor	63
4.2.1. Clasificación de los métodos para detectar el rostro	63
4.2.2. Localización del rostro	65
4.3. Detección de los ojos	67
4.3.1. Trabajos previos para detectar los ojos	67
4.3.2. Inconvenientes para detectar los ojos	67
4.3.3. Método propuesto para detectar los ojos basado en el algoritmo EM	68
4.4. Enfoque bayesiano para el seguimiento del rostro y los ojos	74
4.4.1. Motivos para el seguimiento	74
4.4.2. Descripción del algoritmo Condensation	75
4.5. Seguimiento del rostro	77
4.5.1. Especificación del modelo de seguimiento del rostro	78
4.5.2. Descripción de las redes neuronales artificiales	79
4.5.3. Entrenamiento de la red neuronal artificial	80
4.5.4. Evaluación de la función de densidad	84
4.5.5. Resultados experimentales del seguimiento del rostro	85
4.6. Seguimiento de los ojos	89
4.6.1. Especificación del modelo de seguimiento de los ojos	89
4.6.2. Correlación de modelos	90
4.6.3. Evaluación de la función de densidad	90
4.6.4. Resultados experimentales del seguimiento de los ojos	91
4.7. Cálculo del índice de somnolencia con SVM	94
4.7.1. Descripción de SVM	95
4.7.2. Filtro de Gabor	96
4.7.3. Representación 2D del ojo a través del filtro de Gabor	97
4.7.4. Extracción de características basada en el filtro de Gabor	98

4.7.5. Entrenamiento de SVM	99
4.7.6. Construcción del índice de somnolencia	103
4.8. Detección de la distracción	107
4.8.1. Orientación del rostro	107
4.8.2. Dirección del rostro usando redes neuronales	108
4.9. Velocidad de procesamiento	109
5. Detección de la Somnolencia en el Infrarrojo Cercano	111
5.1. Descripción del sistema en el infrarrojo cercano: diseño y alcance	111
5.2. Detección de los ojos	114
5.2.1. La imagen diferencia	115
5.2.2. Método propuesto para umbralizar de la imagen diferencia	116
5.2.3. Localización de los candidatos a ojos basada en la imagen de bordes	118
5.2.4. Método propuesto para umbralizar la imagen de bordes	118
5.2.5. Verificación y localización de los ojos	120
5.2.6. Eliminación basada en restricciones geométricas	121
5.2.7. Eliminación basada en el clasificador SVM	122
5.2.8. Resultados experimentales en la localización de los ojos	127
5.3. Detección del rostro	127
5.3.1. Modificación del contraste	129
5.3.2. Funciones de proyección y varianza vertical	129
5.3.3. Estimación del rostro a partir de la posición de los ojos	132
5.4. Enfoque bayesiano para el seguimiento de los ojos y el rostro	133
5.4.1. Seguimiento de los ojos	134
5.4.2. Seguimiento del rostro	138
5.5. Cálculo del índice de somnolencia	142
5.6. Detección de la distracción	145
5.6.1. Estimación de la orientación del rostro basada en la posición de los ojos	145
5.7. Velocidad de procesamiento	146

6. Conclusiones y Futuros Desarrollos	147
6.1. Conclusiones	147
6.1.1. En los dos dominios	148
6.1.2. Espectro visible (conducción diurna)	149
6.1.3. Infrarrojo cercano (conducción nocturna)	150
6.2. Futuros desarrollos	152
 A. Plataforma experimental Ivvi	 155
A.1. Sistema de detección de vehículos	158
A.2. Sistema de detección del carril	159
A.3. Sistema de detección de peatones	160
A.4. Sistema de detección de señales de tráfico	161
A.5. Sistema de detección de la somnolencia	162
 B. Diseño y Construcción del Sistema de Percepción	 165
B.1. Sistema de visión en el infrarrojo cercano	165
B.1.1. Elementos del sistema de visión	165
B.1.2. Sistema de visión	166
B.1.3. Sistema de iluminación	168
B.1.4. Sistema de sincronización	175
B.1.5. Integración de los sistemas de visión, iluminación y sincronización .	178
B.1.6. Montaje y resultados	191
 C. Conceptos de Estadística y Probabilidad	 195
C.1. Variables aleatorias, espacio muestral, función de distribución acumulada .	195
C.1.1. Función de densidad de probabilidad	196
C.1.2. Valor esperado y varianza	196
C.1.3. Distribuciones discretas	197
C.1.4. Distribuciones continuas	197
C.1.5. Densidad normal multivariante	198
C.2. Introducción al Algoritmo EM	198
C.2.1. Descripción del algoritmo EM	199

C.2.2. Estimación del modelo paramétrico del color de la piel	200
Bibliografía	203

Índice de figuras

1.1. Accidente causado por los efectos de la somnolencia	1
1.2. Comparación entre el transporte no motorizado y motorizado	6
1.3. Uso de la bicicleta	6
1.4. Producción mundial de vehículos	7
1.5. Efectos de la masificación vehicular	8
1.6. Datos históricos de los accidentes de tráfico	10
1.7. Clasificación de los accidentes de tráfico	14
1.8. Distracción del conductor	17
1.9. Apreciación global de la estructura del sistema de detección de la somnolencia.	18
1.10. Vehículo experimental IvvI.	19
1.11. Estructura del trabajo de investigación.	21
2.1. Clasificación de los sistemas de detección de la somnolencia	24
2.2. Patrón de conducción a partir de la posición del vehículo	25
2.3. Análisis del rostro con una cámara	26
2.4. Electroencefalograma	27
2.5. Sistema de Jiangwei	30
2.6. Sistema de Branzan	31
2.7. Sistema desarrollado por Hayami	32
2.8. Sistema de iluminación diseñado por Ji	33
2.9. Diferencia de imágenes del campo par e impar	33

2.10. Diagrama de detección de los ojos	33
2.11. Resultados de detección de los ojos utilizado por Ji	34
2.12. Copilot	37
2.13. Driver Fatigue Monitoring	37
2.14. Sistema de Seeing Machine	38
2.15. Simulador desarrollado por Bosh	39
2.16. Control de alerta de Volvo	39
2.17. Sistema experimental de Mercedes-Benz	40
2.18. Proyecto AWAKE	41
2.19. Sistema de Mitsubishi	42
2.20. Sistema de SafeTrac	42
2.21. Asiento que detecta el sueño	43
3.1. Apreciación global del sistema de detección de la somnolencia.	48
3.2. Imágenes capturadas con el sistema de visión del vehículo IvvI	49
3.3. Módulos del sistema de visión con iluminación infrarroja.	51
3.4. Espectro electromagnético.	52
3.5. Sistema de visión con iluminación infrarroja	55
3.6. Esquema del sistema de visión con iluminación infrarroja.	56
3.7. Sistema de visión con iluminación infrarroja y su electrónica asociada	57
3.8. Sistema de visión con iluminación infrarroja instalado sobre el vehículo IvvI	58
3.9. Imágenes capturadas con el sistema de visión de IvvI	59
4.1. Esquema del algoritmo en el espectro visible	62
4.2. Información de la región de interés	65
4.3. Resultados del método propuesto para localizar el rostro y los ojos	66
4.4. Ejemplos donde el algoritmo de Viola & Jones falla	66
4.5. Anatomía del ojo	68
4.6. Muestra aleatoria de la forma del ojo	69
4.7. Rectángulos preliminares para detectar los ojos	70
4.8. Resultados de la segmentación de los ojos	72

ÍNDICE DE FIGURAS

4.9. Imágenes utilizadas para localizar la posición de los ojos	72
4.10. Resultado del algoritmo <i>EM</i> al modelar la forma del ojo	73
4.11. Ejemplos donde se localiza la posición de los ojos	73
4.12. El rostro del conductor no es detectado por el método de Viola & Jones en varios instantes de tiempo.	74
4.13. Comparación entre la posición exacta y la posición generada por Viola & Jones	77
4.14. Transformación al espacio imagen	80
4.15. Corrección de la iluminación con el factor gamma y ecualización del histo- grama	82
4.16. Eliminación de las esquinas y operación AND	82
4.17. Ejemplos de la base de datos de rostros	83
4.18. Curva ROC generada sobre la base de datos de prueba de rostros.	84
4.19. Representación de las partículas del algoritmo Condensation sobre el rostro	85
4.20. Probabilidad generada por Condensation en el seguimiento del rostro	86
4.21. Corrección para detectar el rostro del conductor	87
4.22. Resultados y errores generados con el algoritmo Condensation	88
4.23. Modelos para la correlación de los ojos en el algoritmo Condensation	91
4.24. Resultados del algoritmo Condensation en el seguimiento de los ojos	92
4.25. Probabilidad generada por Condensation en el seguimiento de los ojos . . .	93
4.26. Error generado por Condensation en el seguimiento de los ojos	93
4.27. Resultados del seguimiento de los ojos del conductor	94
4.28. Esquema de consulta a SVM.	95
4.29. Representación de la optimización desarrollada en SVM.	96
4.30. Representación del wavelet de Gabor	97
4.31. Imágenes generadas con el filtro Gabor	99
4.32. Ventanas para extraer el vector de características de la imagen del ojo. . . .	100
4.33. Ejemplos de la base de datos de ojos.	100
4.34. Función RBF en dos dimensiones.	101
4.35. Resultado del análisis instantáneo del estado del conductor.	102

4.36. Resultado del análisis del estado del conductor sobre el tiempo	103
4.37. Evolución del índice de Horng <i>et al.</i> y Dong-Wu	104
4.38. Evolución de Perclos	105
4.39. Resultado del análisis del estado del conductor sobre el tiempo	106
4.40. Estimación de la orientación del rostro a partir de la posición de los ojos. .	107
4.41. Evolución de la orientación del rostro a partir de la posición de los ojos. .	108
4.42. Resultados del análisis de la distracción del conductor	109
5.1. Esquema del algoritmo de detección de la somnolencia en el infrarrojo cercano.	112
5.2. Imágenes generadas con el sistema de visión bajo iluminación infrarroja . .	113
5.3. Estructura de la imagen capturada con el sistema de visión con iluminación infrarroja.	113
5.4. Diagrama para detectar los ojos en el infrarrojo cercano.	114
5.5. Separación de la imagen en el campo par e impar	115
5.6. Diferencia de imágenes y su umbralización	117
5.7. Representación del filtro de la mediana.	118
5.8. Imagen de bordes de la imagen diferencia	119
5.9. Estimación de la región de interés del ojo a partir del área de la pupila . . .	123
5.10. Filtro de Gabor aplicado a las imágenes de entrenamiento de SVM	124
5.11. Resultados de PCA de los ojos	125
5.12. Representación de los eigenvalores del PCA de los ojos	126
5.13. Proyecciones en el espacio Eigenrostro	126
5.14. Ejemplos de la detección de los ojos en el infrarrojo cercano	127
5.15. Imágenes generadas con el sistema de iluminación infrarroja instalado sobre el vehículo IvvI	128
5.16. Modificación del contraste	130
5.17. Función de proyección vertical	131
5.18. Resultados obtenidos en la detección del rostro y los ojos.	131
5.19. Modelo paramétrico del rostro a partir de la posición de los ojos.	132
5.20. Resultados de la estimación del rostro a partir de la posición de los ojos . .	133

ÍNDICE DE FIGURAS

5.21. Función de densidad triangular para evaluar los pesos del algoritmo Con-	
densation	135
5.22. Resultados del algoritmo Condensation en el seguimiento los ojos bajo ilu-	
minación infrarroja	137
5.23. Ejemplo donde los ojos se encuentran cerrados	138
5.24. Ejemplos de rostros de la base de datos de entrenamiento	139
5.25. Curvas ROC generadas en la clasificación de los rostros	140
5.26. Probabilidad de las observaciones generadas con el algoritmo Condensation	
en el seguimiento del rostro.	140
5.27. Seguimiento del rostro con el algoritmo Condensation	141
5.28. Respuesta instantánea del sistema propuesto bajo iluminación infrarroja . .	143
5.29. Evolución de Perclos en conducción nocturna	144
5.30. Evolución de la orientación del rostro	145
A.1. Plataforma experimental Ivvi	156
A.2. Arquitectura de los sistemas que componen la plataforma Ivvi.	157
A.3. Interior de Ivvi	157
A.4. Cámaras interiores de Ivvi	158
A.5. Exterior de Ivvi	158
A.6. Detección de vehículos utilizando modelos geométricos	159
A.7. Detección del carril de la carretera	160
A.8. Detección de peatones en el día.	161
A.9. Detección de peatones en la noche	161
A.10. Detección de las señales de tráfico.	162
A.11. Sistema de detección de la somnolencia	163
B.1. Cámara GC400CP	166
B.2. Filtro IR del sistema de visión	167
B.3. Imágenes tomadas con y sin filtro IR	168
B.4. Ley de Snell.	169
B.5. Diseño genérico del sistema de visión con iluminación infrarroja.	170

B.6. Diagrama de radiación del SFH 487	171
B.7. Led tipo jumbo ELJ-810-248 B.	172
B.8. Fuente de alimentación genérica	173
B.9. Fuente de alimentación de los jumbos y de los leds SFH 487	174
B.10. Señal de video compuesto vista desde el osciloscopio.	176
B.11. Integrado LM1881N.	177
B.12. Diagrama del circuito integrado LM1881N	178
B.13. Señales proporcionadas por el LM1881N	179
B.14. Esquema del sistema de visión con iluminación infrarroja.	180
B.15. Conexión de los sistemas de visión y sincronización.	181
B.16. Señal de video compuesto proporcionado por el sistema de visión.	181
B.17. Señal de sincronismo obtenida en el LM1881N.	182
B.18. Conexión del sistema de sincronización-iluminación	184
B.19. Señal de control de potencia de los jumbos.	185
B.20. Comportamiento dinámico de los jumbos	186
B.21. Comportamiento dinámico fuente de corriente jumbos.	187
B.22. Conexión del sistema de sincronización-iluminación	188
B.23. Señales de control de disparo y potencia de los jumbos y del anillo central .	189
B.24. Comportamiento dinámico del grupo de leds SFH 487.	190
B.25. Sistema de visión con iluminación infrarroja.	191
B.26. Esquema de medida de la intensidad radiada.	192
B.27. Sensibilidad BPW34 y área sensible.	192
B.28. Potencia recibida por área de superficie en función de la distancia.	193
C.1. Gráfico de contorno del modelo del color de piel obtenido con el algoritmo EM.	201
C.2. Resultados del modelo de color de piel	202

Índice de tablas

4.1. Resultados del seguimiento del rostro con el algoritmo Condensation	87
4.2. Resultados del seguimiento de los ojos generado por el algoritmo Condensation	94
4.3. Resultados del análisis del estado de los ojos usando <i>SVM</i>	102
4.4. Tiempos de cómputo en cada fase del algoritmo en el espectro visible	110
5.1. Resultados del seguimiento de los ojos con el algoritmo Condensation bajo iluminación infrarroja	136
5.2. Resultados del seguimiento del rostro con el algoritmo Condensation bajo iluminación infrarroja	142
5.3. Resultados del análisis del estado de los ojos bajo iluminación infrarroja . .	143
5.4. Tiempos de cómputo en cada fase del algoritmo en el infrarrojo cercano . .	146
B.1. Características de la cámara GC400CP.	167
B.2. Comportamiento del sistema de iluminación para varias tensiones.	185
B.3. Voltajes teóricos de los jumbos.	185
B.4. Comportamiento del sistema de los leds para varias tensiones.	190
B.5. Potencia recibida en función de la distancia.	193
C.1. Resultados numéricos del algoritmo EM al ajustar el modelo del color de piel sobre el espacio dicromático HS.	201

Introducción

1.1. Motivación

La excesiva cantidad de accidentes de tráfico, a causa de la somnolencia, constituye el principal motivo que ha dado lugar al desarrollo de esta tesis doctoral, ver figura 1.1. Este trabajo se enmarca en los denominados *Sistemas Avanzados de Ayuda a la Conducción* (SAAC) que constituyen la nueva generación de sistemas de seguridad que involucran a las nuevas tecnologías, que son parte de lo que actualmente se denomina *seguridad activa* dentro del campo de los llamados *Sistemas Inteligentes de Transporte* (SIT) [87], [80], [47], [14].



Figura 1.1: Efectos de la somnolencia en el conductor.

Los SIT constituyen un conjunto de tecnologías de la información aplicadas a mejorar la eficiencia, la planificación y la seguridad de todas las áreas del sector transporte, especialmente, infraestructuras y vehículos. En este sentido, la seguridad activa está desarrollada para garantizar el funcionamiento de un vehículo en movimiento y responder a

las ordenes del conductor y en la medida de lo posible prevenir accidentes [14], [50]. En el caso de que un accidente no se pueda evitar, varios programas actúan para minimizar sus efectos, especialmente, sobre los pasajeros.

Por su parte, los SAAC son sistemas que contribuyen a incrementar la seguridad activa y mejorar la eficiencia y confort en la conducción [50], [97], [80], interactuando con el conductor, el vehículo y la infraestructura [67]. Se puede decir que el objetivo de los SAAC es prevenir y contribuir en la reducción de los accidentes de tráfico y constituyen la respuesta tecnológica a un problema que genera miles de tragedias a nivel mundial, muchas de ellas relacionadas directamente con el conductor.

En este aspecto, el conductor siempre se encuentra expuesto a cometer distintos errores voluntarios y/o involuntarios que finalizan en muchos casos en accidentes. Los errores involuntarios se producen por varios motivos, uno de los principales es la somnolencia.

La somnolencia produce la disminución de la capacidad de vigilia afectando negativamente a los sentidos, por ejemplo, mermando la capacidad de percepción, reconocimiento y control del vehículo [124], y a la vez, incrementado la probabilidad de sufrir algún accidente o un sobresalto en la carretera.

La somnolencia y la conducción forman una combinación peligrosa [52] que produce muchos daños tanto a nivel individual (lesiones y muertes de conductores, peatones e incluso pasajeros en el caso de empresas de transporte) como a nivel colectivo (altos costes económicos, saturación de los sistemas sanitarios), generando serios problemas a nivel social.

El estudio de los problemas derivados de la somnolencia (también se incluye a la distracción) en el proceso de conducción es un tema que cada vez atrae la atención de más investigadores provenientes de diferentes áreas, como son: medicina, ingeniería, psicología, etc.; que procuran desarrollar nuevos métodos para construir sistemas que sean capaces de prevenir y disminuir los accidentes producidos por este percance.

El problema no es fácil de resolver debido a que el mundo contemporáneo, donde cada vez el uso de los vehículos va en aumento, está lleno de factores que contribuyen al

1.2 Somnolencia y distracción

incremento de los accidentes de tráfico, entre ellos se pueden citar: las distancias entre el trabajo y el hogar cada vez son mayores, el incremento del ruido en las ciudades desencadena el aumento del estrés, la contaminación del aire que contribuye en el aumento de enfermedades respiratorias y alergias. A estos factores externos que intervienen en las acciones que involucran al tráfico hay que añadirles los nuevos provenientes de los accesorios incorporados en los vehículos (sistemas de posicionamiento global, radios, teléfonos móviles, etc.) que originan distracción en el conductor, contribuyendo negativamente en las estadísticas de tráfico.

Para finalizar esta sección, es preciso destacar que la mayoría de las empresas del sector automotriz han encontrado en el tema de la seguridad una nueva vía para diferenciarse frente a la competencia; es así que sus departamentos de investigación, desarrollo e innovación (*I+D+i*) se encuentran trabajando activamente en la construcción de distintos modelos y tipos de SAAC, por ejemplo, sistemas de detección de peatones, sistemas de detección de señales de tráfico, sistemas de control de la velocidad y por supuesto, sistemas de vigilancia del conductor, entre otros. Esto demuestra que el tema de la seguridad está en auge y que aún resta mucho trabajo hasta llegar a alcanzar resultados óptimos que cumplan estándares de calidad altamente fiables.

1.2. Somnolencia y distracción

En este apartado se describen los conceptos de somnolencia y distracción, luego, sus características y efectos.

1.2.1. Definición de somnolencia

Según la Real Academia Española [46],

“la somnolencia es la sensación de pesadez y torpeza de los sentidos motivadas por el sueño”

concepto que concuerda plenamente con lo referido anteriormente en la introducción y descrito por Ji *et. al.* [124], y Knipling [130].

La definición es claramente contraria a las recomendaciones que se deben cumplir para conducir, es decir, estar descansado y completamente alerta ante cualquier acontecimiento.

1.2.2. Definición de distracción

Al igual que en el caso anterior, la definición se tomó de la Real Academia Española [46]:

“distracción es la acción y efecto de distraer”

utilizando esta definición se puede decir que la distracción es el efecto de apartar la atención de alguien del objeto, también se la define como distancia y separación.

1.2.3. Factores que producen somnolencia y sus consecuencias

La somnolencia se presenta por muchos factores entre los cuales se encuentran: mucho tiempo al volante, aburrimiento, consumo de alcohol y/o drogas, enfermedades y desórdenes del sueño (narcolepsia, etc.), consumo de medicamentos que producen sueño, etc.

Las principales consecuencias de su presencia son:

- Perjudica el tiempo de reacción
- Disminuye la capacidad de actuación, de concentración y de vigilancia
- Incrementa el comportamiento agresivo y el malhumor, etc.

1.2.4. Características visuales que indican la presencia de somnolencia

Las personas en estado de somnolencia exhiben ciertas características faciales que se pueden reconocer sobre el rostro, es así que, la cabeza, los ojos y la boca brindan suficiente información visual para conocer cuando un conductor presenta somnolencia.

Entre las principales características que determinan la presencia de somnolencia se encuentran: dificultad para enfocar la vista (mirada) y/o mantener los ojos abiertos, bostezar repetidamente, fruncir la ceja, dificultad para mantener la cabeza erguida y dificultad para enfocar y mantener la mirada estable.

1.3 La industria automotriz, su desarrollo, evolución y consecuencias

Otra fuente de información visual es la constituida por el rostro y la cabeza, los mismos que indican el nivel de atención, la dirección de la mirada y el nivel de fatiga. La orientación ideal del rostro es frontal cuando el conductor presta la suficiente atención a la conducción, cualquier otra orientación por períodos extendidos de tiempo o por cambios frecuentes indican fatiga o falta de atención. En el mismo contexto, la posición del rostro es otro indicador de la fatiga y la falta de atención, tal como lo anotan en sus trabajos Ji *et al.* [124] y Bergasa *et al.* [78].

Estas características visuales son aprovechadas por distintos investigadores para desarrollar sistemas de vigilancia del conductor utilizando análisis de imágenes y técnicas de Inteligencia Artificial.

Por los motivos de esta tesis solo se han expuesto las características visuales que indican somnolencia, pero se debe anotar que existen otras características para detectarla, por ejemplo, la alteración del ritmo cardiaco y de las ondas cerebrales [12], [143], [135], [8], [26].

1.3. La industria automotriz, su desarrollo, evolución y consecuencias

La industria automotriz, desde sus inicios a finales del siglo *XIX* (en 1.886 Carl Benz registra la primera patente de automovil), ha ido evolucionando rápidamente por la necesidad que tiene el hombre de contar con medios de transporte cada vez más eficaces para desarrollar sus distintas actividades, llegado a convertirse en una industria pionera en varios campos, entre ellos, los que están relacionados con el desarrollo de la tecnología aplicada al confort, la comodidad y la seguridad.

Actualmente, no cabe duda que los vehículos, en general grandes y pequeños, son necesarios para desarrollar la mayoría de las actividades de la vida moderna, es así, que su uso se ha ido masificando y hoy en día se encuentran al alcance de la mayoría de las personas (en especial en los países desarrollados y cada vez más en los países emergentes). El fenómeno de la masificación automotriz, que se puede observar en el gráfico de barras de

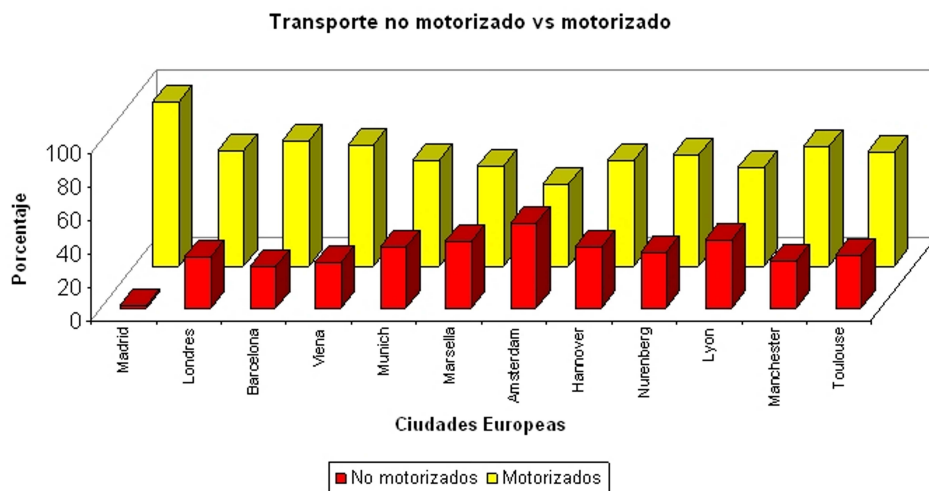


Figura 1.2: Comparación entre el transporte no motorizado y motorizado en varias ciudades europeas (Fuente: OICA, 2006).

la figura 1.2, muestra que en el caso de varias ciudades europeas el transporte motorizado se encuentra por encima del transporte no motorizado, por ejemplo, el uso de bicicletas; que en el caso de los Países Bajos es usada masivamente y se presenta como una alternativa de transporte. La figura 1.3 exhibe la imagen de un típico día en Holanda donde la bicicleta es la protagonista.



Figura 1.3: Uso de la bicicleta en la ciudad de Amsterdam, Holanda.

1.3 La industria automotriz, su desarrollo, evolución y consecuencias

1.3.1. Producción mundial de vehículos

La producción mundial de vehículos en el año 2.007 alcanzó la cantidad de 73 millones de unidades [111], y según la tendencia calculada sobre los registros históricos se espera que la producción para los próximos años seguirá en aumento, debido a que se ha mantenido al alza desde el año 2.001. A todo esto, varias proyecciones indican que los mercados emergentes, especialmente los de China, Rusia e India contribuirán a incrementar este crecimiento porque su demanda va en aumento [112]. Para corroborar esta información se presenta la figura 1.4, donde se exhibe la producción mundial de automóviles desde el 2.001 hasta el 2.007.



Figura 1.4: Datos históricos de la producción mundial de vehículos (en millones) (Fuente: OICA, 2008).

Sin embargo, debido a los problemas generados por la crisis financiera internacional, desatada a finales de 2.008, se podría esperar que estas cifras varíen y su tendencia vaya a la baja o se mantenga constante durante los próximos años.

1.3.2. Congestión vehicular y contaminación del aire

Si bien es cierto, la industria automotriz ha generado y seguirá generando miles de puestos de trabajo y mejorando la calidad de vida de las personas; también es cierto que ha contribuido en el desarrollo de problemas de gran escala; entre ellos se pueden citar, la congestión vehicular y la contaminación del aire. La congestión se produce porque existe un límite físico de las infraestructuras y cuya capacidad está llegando a su punto de saturación y ya no es capaz de soportar la creciente cantidad de vehículos. Por otro lado, la contaminación del aire, especialmente en las grandes ciudades, se debe en gran parte al uso excesivo de los vehículos que generan gases que contaminan el ambiente.



(a)



(b)

Figura 1.5: Efectos de la masificación vehicular: (a) congestión vehicular en distintas autopistas, (b) contaminación del aire en la ciudad.

1.4. Accidentes de tráfico

El estilo de vida contemporáneo se encuentra ligado a la velocidad y a la rapidez, a un estilo de vida que exige resultados (eficaces) en breves períodos de tiempo, lo que Hagemeyer [82] denomina la sociedad de “24-horas”. A la par de esta tendencia, el crecimiento de las ciudades sigue en aumento, cada vez, la gente se congrega en las grandes urbes donde se concentran el trabajo y las actividades de la vida moderna. Según las Naciones Unidas se estima que para el año 2.010 el 51,3 % de la población mundial será urbana [25].

La necesidad de movilizarse, en especial en las grandes ciudades, ha generado que se

1.4 Accidentes de tráfico

incremente la demanda de vehículos en todas sus versiones (camiones, furgones, turismos, motos, autobuses, etc.) con el consiguiente aumento de la contaminación y la congestión; estos factores, entre otros, generan estrés y cansancio en la población en general, y en particular en los conductores; dichos factores influyen directamente en el incremento de los accidentes de tráfico (Hagenmeyer [82] en sus conclusiones hace esta afirmación para el caso de la sociedad occidental).

El problema alcanza tal magnitud y es tan grave que incluso se ha llegado a decir que los accidentes de tráfico son la lacra de la sociedad moderna [24], porque han llegado a constituir en la actualidad *la novena causa de muerte en el mundo y en la principal causa de traumatismos que ocasionan la pérdida de la salud*; y para el año 2.030 se pronostica que se convertirá en la *tercera causa de muertes en el mundo* [150]. Teniendo esto en cuenta, actualmente el conducir se ha convertido en la actividad más peligrosa dentro del ambiente laboral [51] y en el mayor problema de salud pública en Europa [69], donde cada año mueren alrededor de 45.000 personas y 1,5 millones sufren heridas por causa de esta siniestralidad. Además se estima que su coste asciende al 2% del producto interno bruto (PIB), añadiendo los costes de las indemnizaciones a las víctimas, el perjuicio bordea los 160.000 millones de euros.

En América cada año se producen alrededor de 130.000 muertos, más de 1,2 millones de personas sufren traumatismos, y miles de ellas resultan discapacitadas como consecuencia de los choques o atropellos en la carretera. Solo en Estados Unidos, su valor asciende a 230.000 millones de dólares, según lo informa la FHWA Office of Safety Integration.

En la revista Lukor, Noticias y Motor [89] se presentan las siguientes estadísticas de accidentes de tráfico, citadas por el director del Instituto Nacional de Tránsito y Seguridad Vial (INTRAS): en el año 2.002 en el mundo se produjeron alrededor de 1,5 millones de muertes por causa de accidentes de tráfico y cerca de 50 millones de heridos, lo que supone que cada día que pasa mueren unas 4.080 personas y hay más de 132.000 heridos. Por su parte, la Organización Mundial de la Salud (OMS) predice que estas cifras seguirán en aumento durante los próximos años principalmente por el incremento de vehículos en las carreteras [35].

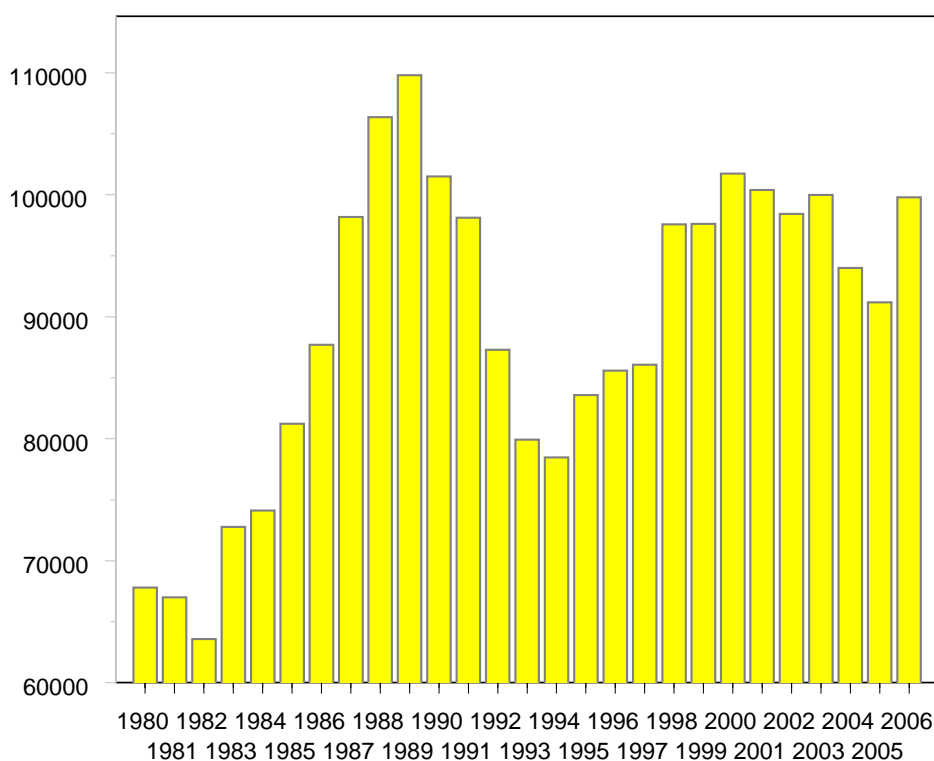


Figura 1.6: Datos históricos de los accidentes de tráfico con víctimas (muertos y heridos) en España (Fuente: DGT, 2.008).

El director del INTRAS [89] también indica que en el caso español se genera alrededor de 15 muertos y más de 450 heridos de cierta gravedad cada día, con costes anuales que ascienden a 16.000 millones de euros, esto significa que cada español invierte 400 euros anuales en pagar esta contingencia. Por este motivo, constituyen *la quinta causa de muerte y producen el 40 % de las minusvalías* [24]. En la figura 1.6 se observa la evolución de los accidentes de tráfico desde 1.980 hasta el 2.005 en España.

1.4.1. División de los accidentes de tráfico

Entre los principales factores que ocasionan los accidentes de tráfico se pueden citar los siguientes:

1.4 Accidentes de tráfico

- **Alcohol** : Es común escuchar que se producen accidentes de tráfico por el consumo de bebidas alcoholicas, en general, los gobiernos hacen esfuerzos por estimular a la población a no consumir alcohol antes de conducir. Pero al parecer estos esfuerzos no tiene la suficiente acogida, es así que en España se estima que 1.500 personas mueren al año y 50.000 resultan heridas en accidentes de tráfico relacionados con el consumo del alcohol, y el 10 % de los siniestros más graves están provocados por el consumo de drogas ilegales [44]. Sin embargo, en el enfoque de este trabajo, esto no es relevante y, de aqui, no será discutido en mayor detalle.
- **Enfermedades** : Un estudio desarrollado en la Unión Europea demuestra que las personas que padecen alguna enfermedad tiene mayor riesgo de estar implicadas en un accidente de tráfico [37], [122]. Los trastornos neurológicos y las enfermedades mentales son las que mayor riesgo de accidentalidad presentan. Otro grupo que contribuye a aumentar los accidentes se encuentran formado por las enfermedades del sueño y las alergias.
- **Distracción** : Varios factores externos e internos al vehículo, a menudo, captan la atención del conductor, distrayéndolo de su labor de conducción y finalizando en un susto grave o un accidente. Liang *et al.* [159] indica que la distracción genera una buena parte de las estadísticas de los accidentes de tráfico.
- **Somnolencia (Fatiga)**: Genera muchos accidentes porque el conductor presenta de manera inoportuna síntomas de sueño, incrementando la posibilidad de sufrir un accidente. Las enfermedades relacionadas con el sueño, como la apnea, también ayudan a incrementar este percance.
- **Meteorología** : Las malas condiciones del ambiente ayudan a incrementar el número de accidentes de tráfico, entre los principales factores está el exceso de neblina y niebla, vientos fuertes y la lluvia, esta última ocasiona que las carreteras se vuelvan inestables.
- **Velocidad** : La velocidad es un factor inherente al desplazamiento de los vehículos, pero que en muchos casos debido a la imprudencia y/o impericia de los conductores termina siendo la causante de un accidente. Es así que, el exceso de velocidad junto

a otros factores como el omitir las señales de tráfico contribuyen negativamente en el aumento de las estadísticas de accidentes de tráfico con muerte en aproximadamente un 30 % [38].

En las próximas secciones se analizarán las estadísticas que generan la somnolencia y como ésta contribuye en el aumento de los accidentes de tráfico.

1.4.2. Accidentes de tráfico por somnolencia

De los factores que se han mencionado en el párrafo anterior, en este trabajo, se toma especial interés a los accidentes causados por la somnolencia, y en menor medida a los accidentes que se producen por la distracción.

A pesar de que la mejor manera de evitar los accidentes es siendo responsable [103], existen situaciones que no dependen del conductor, la somnolencia es una de ellas; reduce en un 68 % la capacidad de percibir las señales de la carretera [132] y la reacción frente al volante es un 74 % más lenta, según informa la Deutsche Verkehrswacht, organización alemana dedicada a la seguridad en las carreteras [146].

A la velocidad de 100 kilómetros por hora el vehículo recorre aproximadamente 28 metros en un segundo, desde esta perspectiva resulta fácil imaginar lo que sucede si disminuye el estado de vigilia durante cortos lapsos de tiempo durante la conducción debido al adormecimiento involuntario generado por la presencia del cansancio.

En la figura 1.7(a) se muestra la división de los accidentes de tráfico, poniendo énfasis en la relación que existe entre los accidentes por somnolencia y por distracción. Fletcher *et al.* [81] señala que estos dos factores generan el 20 % del total de accidentes, aunque otros autores le atribuyen un porcentaje mayor.

Los investigadores Knippling y Wierwille [130] y las referencias electrónicas [45] (Trastornos del sueño) y [39] (Diagnóstico y tratamiento a pacientes con problemas para dormir) indican que los accidentes de tráfico por somnolencia ocurren frecuentemente en horas de la noche en especial en el período comprendido entre las 11 p.m. y las 6 a.m. seguido del período comprendido en las primeras horas de la tarde, esto es, entre las 14 p.m. y las 16

1.4 Accidentes de tráfico

p.m. Sin embargo, no hay que pasar por alto otros momentos porque la somnolencia se puede presentar a cualquier hora, especialmente, en personas que han consumido algún tipo de medicamento que produce sueño o por haber tomado bebidas alcohólicas.

Utilizando los informes del General Estimates System (GES) del NHTSA [109] se tiene la siguiente información: la mayoría de accidentes de esta clase se producen en autopistas que tienen establecido el límite de velocidad entre 88 y 104 kilómetros por hora, el 85 % de estos accidentes lo producen vehículos del tipo turismo¹ y, en la mayoría de casos, el accidente sucede en tramos rectos de la vía (83 % en tramos rectos y 17 % en curvas), figura 1.7(c). En el 78 % de los casos de accidentes el conductor es el único ocupante del vehículo, figura 1.7(d).

En lo que se refiere al género y la edad, los accidentes generados por la somnolencia se encuentran fuertemente relacionados; es así que los hombres cometieron 1,8 veces más accidentes que las mujeres, ver la figura 1.7(f) . Además, los conductores menores a treinta años son los causantes del 62 % de esta clase de percances, en la figura 1.7(e) se presenta gráficamente lo anotado.

Por otra parte, ampliando el estudio al caso de continentes se tiene que en Europa, los accidentes de tráfico producto de la somnolencia se ubican entre el 10 % y el 20 % del total de ellos, lo que genera cuantías comprendidas entre 16.000 y 32.000 millones de euros. El alto coste en vidas y recursos económicos ha hecho que se consolide un acuerdo entre las autoridades europeas, con la finalidad de reducir las cifras antes mencionadas, fijándose como objetivo reducir a la mitad la cifra de accidentes de tráfico para el año 2010 [105].

Si se toma en cuenta el porcentaje mencionado para el caso de la UE en conjunto con las cifras del INTRAS [89], significa que en España se producen entre 1,5 y 3 muertos y más de 45 heridos, al día, por causa de la somnolencia. En cifras económicas significa que los accidentes por somnolencia cuestan entre 1.600 y 3.200 millones de euros. A este análisis hay que añadir que España tiene el índice más alto de accidentes de tráfico de toda la UE, con lo cual las cantidades arriba mencionadas tienden a incrementarse.

¹En la página electrónica de la DGT se identifican los siguientes tipos de vehículos: (i) Camiones y furgones, (ii) Autobuses, (iii) Turismos, (iv) Motocicletas, (v), Tractores industriales y (vi) Otros.

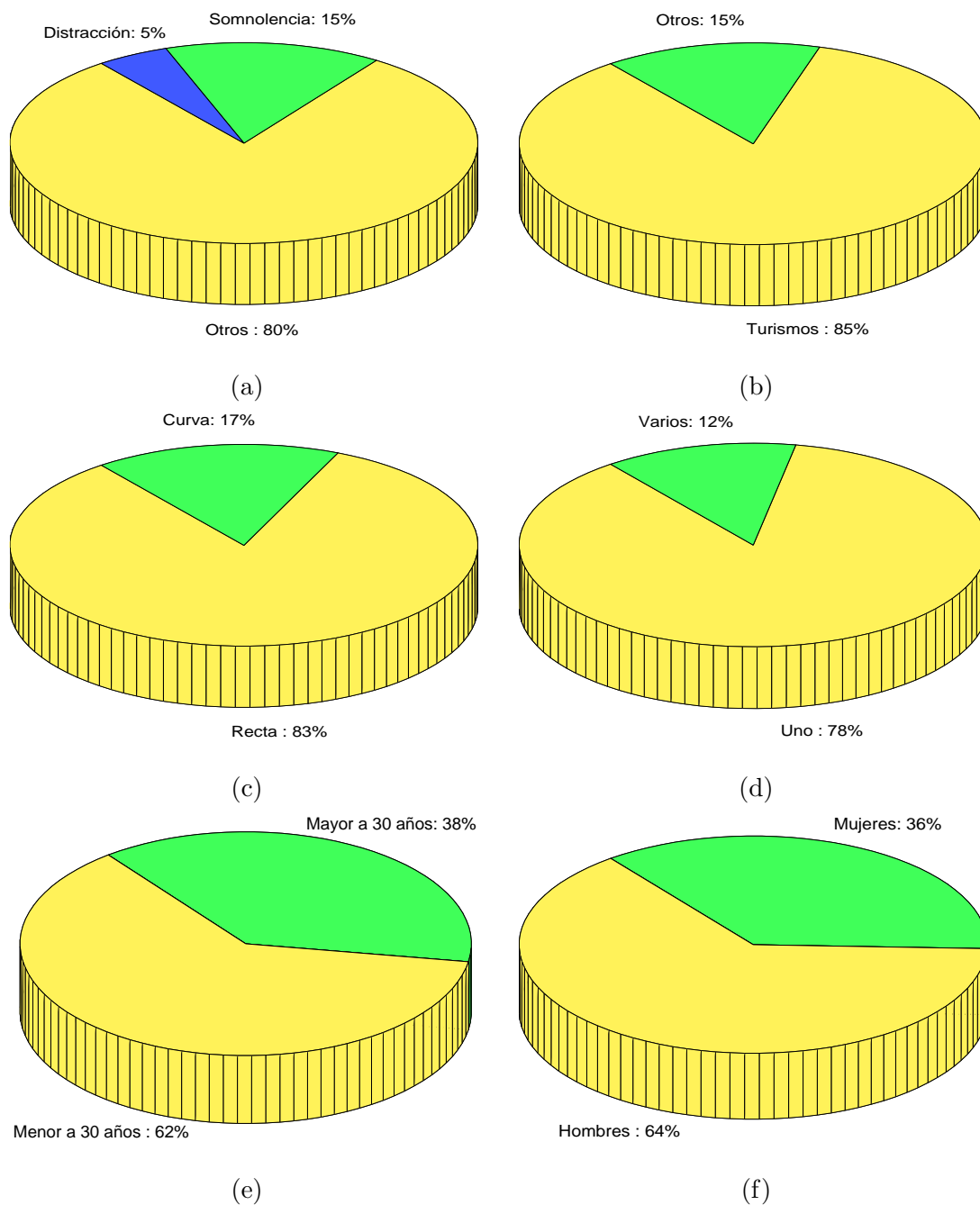


Figura 1.7: Accidentes de tráfico: (a) división basada en la somnolencia y la distracción, y varios gráficos que indican los accidentes de tráfico generados por la somnolencia en el siguiente orden, (b) tipo de vehículo, (c) tipo de carretera (d) número de ocupantes, (e) edad del conductor, (f) género.

1.4 Accidentes de tráfico

En otro estudio relacionado al caso español, desarrollado por la *Sociedad Española de Neumología y Cirugía Torácica* (SEPAR) [149], se concluye que alrededor de 40.000 accidentes de tráfico son producto del sueño. El estudio también señala que las personas que conducen con sueño tienen hasta 13 veces más riesgo de sufrir algún accidente. También los investigadores del estudio expresan:

*“Esto significa –explica Masa– que la somnolencia se configura, junto con el alcohol, como uno de los factores más importantes a la hora de explicar buena parte de los accidentes de tráfico atribuibles al llamado **fallo humano**, que por otra parte son la inmensa mayoría”.*

Para el caso francés se tiene que uno de cada tres accidentes de tráfico son producidos por la somnolencia según los informes del año 2.003 (Federation of French motorway and toll facility companies 2.006) [110], esta información es ratificada por Dong and Wu [152] al indicar que el 20,6 % de las muertes y el 14,6 % de los heridos se generan por los síntomas de la somnolencia.

Según un estudio desarrollado en Inglaterra se estima que el 20 % de todos los accidentes de tráfico en autopistas son producto de la somnolencia, mientras que el 16 % de los mismos se produce en la carretera [82]. Por otra parte, una investigación publicada por el German Highways Department [147] menciona que uno de cada cinco accidentes de tráfico está causado por conducir cansado. La Deutsche Verkehrswacht [146] informa que dormirse al volante y las distracciones son las responsables de uno de cada cuatro accidentes en las autopistas alemanas.

Longhurst [58] también aporta información de lo que sucede en Australia, económicamente, los accidentes por somnolencia implican aproximadamente 1.500 millones de dólares americanos. Branzan *et al.* [1] por otra parte, presenta información de lo que sucede en Canadá, en donde el 20 % de los conductores (aproximadamente 4,1 millones) han sido víctimas de la somnolencia al menos una vez en los últimos 12 meses.

En Estados Unidos, a causa de la somnolencia se han producido alrededor de 100.000 accidentes al año, obviamente, ellos han dejado como saldo, la impresionante cantidad de

150.000 muertos y 71.000 heridos. Pasando a cifras económicas, se menciona que los costes alcanzan aproximadamente la cantidad de 12.500 millones de dólares al año [88], [12].

En el campo de la aviación se estima que el 21 % de los incidentes producidos en USA se deben a la presencia de la somnolencia, el estudio fue desarrollado por la NASA Aviation Safety Reporting System [82]. En el mismo escenario, en Australia 7 % de los accidentes de aviación tienen relación con la fatiga [69], [82]. Curiosamente, el incidente ocurrido en 1.997 entre la estación espacial MIR y la nave espacial Progress 234 fue causada por problemas de somnolencia del piloto de la nave [69].

1.4.3. Consejos para evitar la somnolencia durante la conducción

Por todas estas razones, el problema de la somnolencia se encuentra latente y puede afectar a cualquier conductor en cualquier instante, a pesar de que tome ciertas medidas como bajar la ventanilla o subir el volumen de la música.

Esto sucede porque la población en general desconoce las consecuencias de la privación del sueño o de la modificación de los ritmos biológicos. Los expertos recomiendan varios consejos para evitar dormirse frente al volante, uno de ellos es detenerse, bajar del vehículo y caminar por unos minutos, hidratarse y tomar bebidas que contengan cafeína.

1.4.4. Somnolencia y distracción como factores de accidentes de tráfico

Los accidentes de tráfico por somnolencia difieren de los accidentes por distracción en esencia porque los primeros se producen por factores externos, y en general, involuntarios, como pueden ser la sensación de sueño, estrés, aburrimiento; mientras que los segundos se producen, en la mayoría de los casos, por actos imprudentes cometidos por los conductores, por ejemplo, hablar por el teléfono móvil, manipular el reproductor de CDs del vehículo, mirar el GPS, o simplemente, conversar con otra persona mientras el vehículo se encuentra en marcha. Esto indica que en la medida que se incorporan elementos al vehículo también se añaden factores de distracción, debido a la cantidad de información que debe ser procesada por el conductor [159], la figura 1.8 muestra un ejemplo de una acción imprudente. Además de lo anterior, existen otras fuentes que generan distracción y

1.5 Objetivos

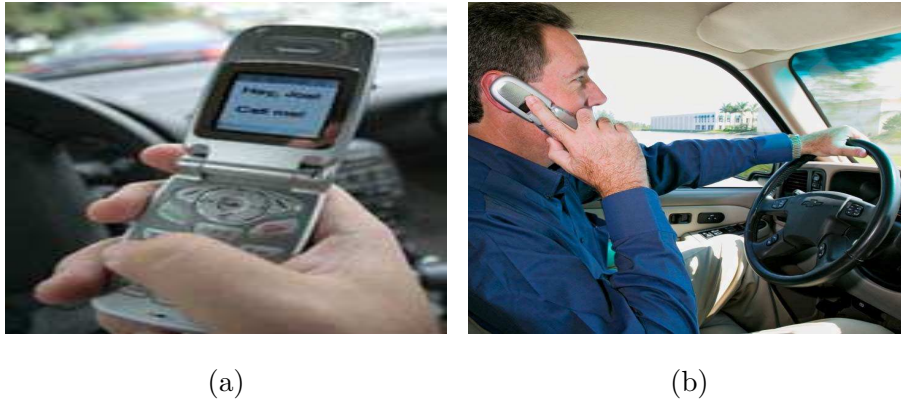


Figura 1.8: Ejemplo de distracción del conductor: el teléfono móvil genera distracción mientras se conduce por dos razones, (a) manipulación y (b) uso.

que no dependen del conductor, entre ellas se tiene el cruce inesperado de peatones sobre la carretera, construcciones mal señalizadas, vehículos de emergencia, etc.

Liang *et al.* [159] indica que entre el 13 % y el 50 % de los accidentes de tráfico son producto de la distracción. Por su parte, Brand *et al.* [139] presenta la siguiente información basado en el informe de accidentes de tráfico del Gobierno de Queensland (Australia), el 31 % de todos los accidentes son producto de la distracción y el 5 % de la fatiga, de donde concluye que el 20 % de todos los accidentes caen en la categoría de fatiga y distracción. En Francia uno de cada cinco accidentes está relacionado con el alcohol, las drogas y la distracción (Federation of French motorway and toll facility companies 2.006).

Curiosamente, las estadísticas de accidentes indican que los meses de verano del hemisferio norte, es decir, julio y agosto, inciden en el incremento de los accidentes por distracción en un 50 % [36].

1.5. Objetivos

El objetivo de esta tesis es *el diseño, desarrollo y construcción de un sistema avanzado de ayuda a la conducción embarcado en un vehículo real, para detectar automáticamente, en tiempo real y bajo condiciones variables de iluminación, la somnolencia del conductor; utilizando para este propósito técnicas de Visión por Computador e Inteligencia Artificial.*

Específicamente, el trabajo se centra en lo siguiente:

- Construir un método robusto para localizar, seguir y analizar los ojos del conductor bajo condiciones variables de iluminación presentes tanto en el día como en la noche.
- Implementar un índice de somnolencia basado en el análisis de los ojos del conductor, lo que permitirá medir su estado de vigilia y evitar que se duerma.
- Implementar un índice para detectar la distracción mediante el análisis de la orientación del rostro y la posición de los ojos, lo que permitirá alertar al conductor si pierde la atención mientras conduce.
- Construir algoritmos eficaces y fiables que sean capaces de procesar la información en tiempo real. Este aspecto está relacionado con tratar de reducir al máximo el número de falsos negativos (cuando no se detecta somnolencia/distracción), y a la vez, reducir la tasa de falsos positivos (cuando se alerta de somnolencia/distracción cuando ésta no existe). Evidentemente, las consecuencias de no detectar somnolencia puede ser mucho más grave que dar una falsa alarma. Sin embargo, los falsos positivos suelen ser molestos por generar falsas alarmas, que a largo plazo molestan al conductor y podrían convertirse en un factor de distracción.
- Instalar el sistema de detección de la somnolencia en un vehículo y desarrollar varios experimentos en condiciones reales de conducción, utilizando para ello la intervención de varios conductores.

1.6. Estructura de la tesis

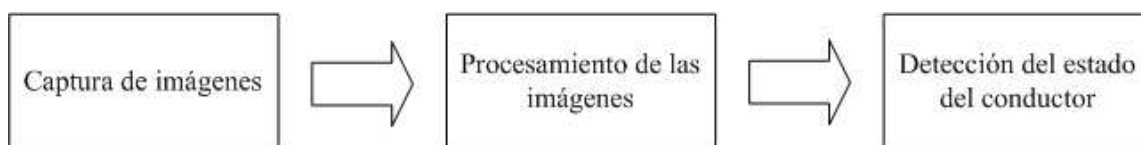


Figura 1.9: Apreciación global de la estructura del sistema de detección de la somnolencia.

La tesis está estructurada de acuerdo al esquema de la figura 1.9 que indica a grosso modo el trabajo a desarrollar; una vez capturada la imagen, el trabajo se centra en extraer

1.6 Estructura de la tesis

la información necesaria del rostro y los ojos del conductor para determinar su estado. Extraer la información es un trabajo complicado y ha motivado a que hubiere que ensayar varias técnicas hasta alcanzar los resultados que satisfagan los objetivos planteados, todo esto dentro de los límites que vienen impuestos por la ciencia y tecnología actuales, es así, que de acuerdo a las condiciones del problema, para extraer la información se lo ha dividido en dos escenarios: en el espectro visible y en el infrarrojo cercano. Esta propuesta permite desarrollar un sistema para el día y uno para la noche.



Figura 1.10: Vehículo experimental IvvI.

Este trabajo no es aislado, es parte de los sistemas que integran el vehículo IvvI, figura 1.10, que consiste en una plataforma experimental diseñada para desarrollar y probar SAAC [90], [70]. En la actualidad existen varios módulos instalados sobre IvvI, los mismos que se mencionarán a continuación: detección de señales de tráfico, detección de peatones, detección de vehículos, detección de las líneas de la carretera y, finalmente, el módulo desarrollado en esta tesis que abarca los escenarios del espectro visible y el infrarrojo cercano.

Una vez que ha sido enmarcada la tesis dentro de una problemática, la estructura del presente documento se compone de varios capítulos, de este modo, la memoria se configura de la siguiente manera:

- i) En el capítulo 2 se presenta el estado del arte, donde se describe detalladamente el estado actual de las técnicas utilizadas para abordar el problema de la somnolencia del conductor. La división del estado del arte se la realiza de acuerdo al tipo de iluminación, es decir, en el espectro visible y en el infrarrojo cercano. Para continuar con un resumen de los desarrollos generados por las empresas privadas, y finalizar con las perspectivas para el futuro en el desarrollo de sistemas de ayuda a la conducción.
- ii) A continuación, en el capítulo 3 se presenta la apreciación global y el sistema de percepción. En la primera parte se introducen los escenarios donde trabajará el sistema y se enumeran los módulos que lo compondrán. En la segunda parte, se describirá el sistema de percepción el diseñado y construido con iluminación infrarroja para resaltar el brillo de la pupila y así facilitar la localización de los ojos durante la noche.
- iii) El capítulo 4 describe el algoritmo desarrollado para la detección de la somnolencia en el espectro visible, esto es, el sistema que trabaja en el día, utilizando imágenes monocromáticas.
- iv) En el capítulo 5 se presenta el algoritmo del sistema de detección de la somnolencia en la noche usando iluminación infrarroja, es decir, se presenta el algoritmo que trabaja en la noche aprovechando las propiedades espectrales de la pupila bajo la influencia de la iluminación infrarroja.
- v) Finalmente, en el capítulo 6 se presentan las conclusiones a las que se ha llegado al realizar esta investigación, también se incluyen los futuros aportes que se pretende seguir desarrollando en este campo del conocimiento.
- vi) En el apéndice se incorporan tres secciones, la primera incluye una descripción detallada del vehículo experimental IvvI, la segunda presenta la información necesaria para construir el sistema de visión con iluminación infrarroja, finalmente, la tercera parte hace referencia a los siguientes temas:

1.6 Estructura de la tesis

- Conceptos de Estadística y Probabilidad
- Introducción al algoritmo EM, con una aplicación para construir un modelo paramétrico del color de la piel

Para cerrar esta sección y finalizar este capítulo, en la figura 1.11 se presenta un gráfico que ilustra por medio de un esquema los capítulos que componen la memoria.

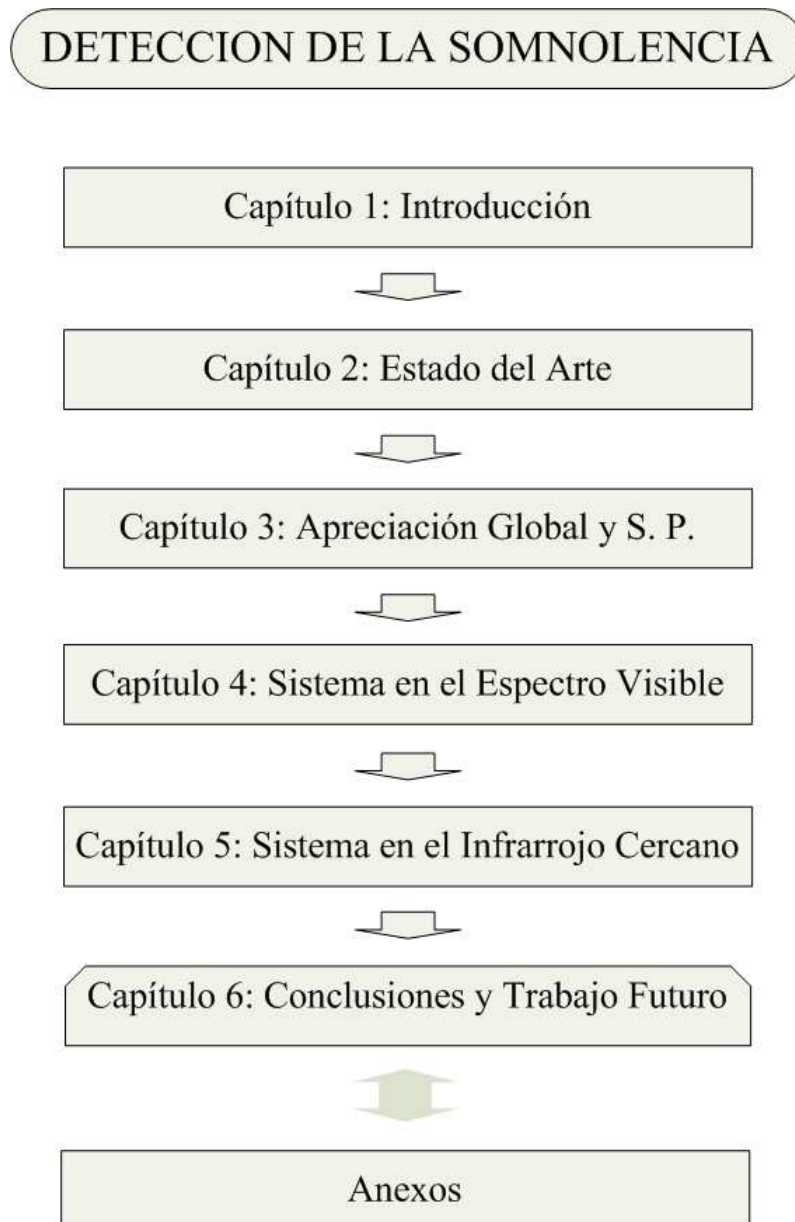


Figura 1.11: Estructura del trabajo de investigación.

Detección de la somnolencia:

Estado del Arte y Perspectivas

Varios e importantes trabajos relacionados con el tema del análisis del conductor han sido publicados en los últimos años dentro del sector automotriz, cada uno de ellos han aportado soluciones parciales al problema global, entre ellos se incluyen principalmente la somnolencia y la distracción, que se encuentran entre los principales factores que causan errores durante el proceso de conducción y en muchos casos finalizan en graves accidentes. Por estos motivos, en este capítulo se presenta el estado del arte en el tema del análisis de la somnolencia del conductor, lo que incluye, la definición del índice más utilizado para medirla, luego se prosigue con la descripción de los sistemas que las empresas privadas han propuesto para abortar este problema, y finalmente se concluye con las perspectivas de su desarrollo, ventajas y comercialización.

2.1. Métodos utilizados para la detección de la somnolencia

Durante los últimos años, motivados por el incremento de los accidentes de tráfico y ayudados por el aumento significativo del poder de cómputo de los ordenadores, se han

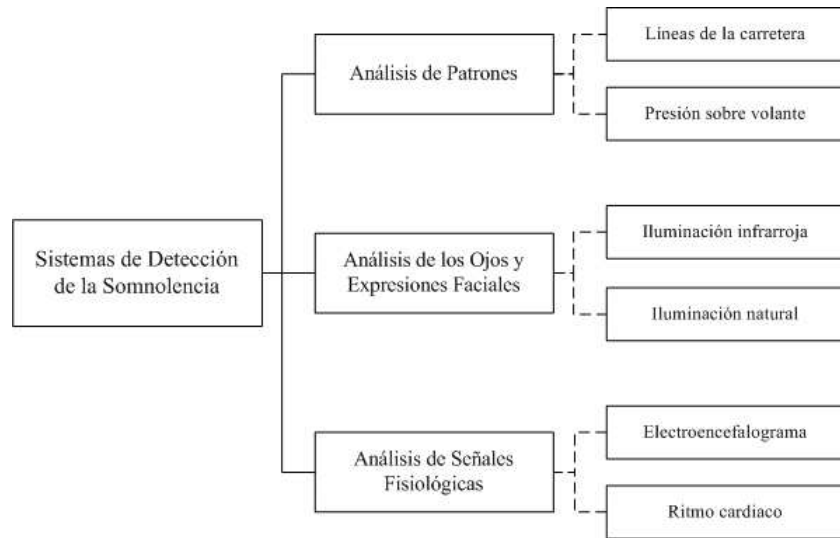


Figura 2.1: Clasificación de los sistemas de detección de la somnolencia del conductor.

presentado varios métodos y propuestas que involucran el uso de distintas tecnologías para enfrentar el grave problema de la somnolencia del conductor. Estos métodos, según Chang *et al.* [12] y Ji *et al.* [125] pueden ser clasificados en tres categorías, enumeradas a continuación e ilustradas en la figura 2.1.

- **Categoría 1:** Análisis basados en patrones de conducción
- **Categoría 2:** Análisis basados en los cambios físicos de los ojos y las expresiones faciales usando procesamiento de imágenes
- **Categoría 3:** Análisis basados en el cambio de las medidas fisiológicas

En las próximas secciones se presentará la información más importante de cada una de estas categorías, previo a un estudio más detallado de la segunda de ellas.

2.1.1. Análisis basados en patrones de conducción

En el primer caso se construyen patrones de conducción generados mediante el desarrollo de diferentes experimentos. Los patrones de conducción se generan a partir de parámetros medibles tanto sobre el conductor y/o sobre el vehículo con su entorno, por ejemplo, la presión de las manos que ejerce el conductor sobre el volante es un patrón de

2.1 Métodos utilizados para la detección de la somnolencia

conducción. Este método no es fácil de modelar, y en muchos casos puede ser necesario diseñar un modelo distinto dependiendo de las características de vehículo y/o conductor. Otros parámetros, tomados sobre el vehículo, que se incluyen en esta clasificación, son la velocidad, la posición lateral, el ángulo de desviación, etc. que han demostrado ser buenos indicadores del nivel de vigilia del conductor, dentro de sus limitaciones [95]. La empresa AssistWare Technology [145] utiliza esta técnica en su propuesta, al igual que el trabajo de Mattsson [77] que construye un patrón basado en la posición del vehículo sobre la carretera, un ejemplo se observa en la figura 2.2.

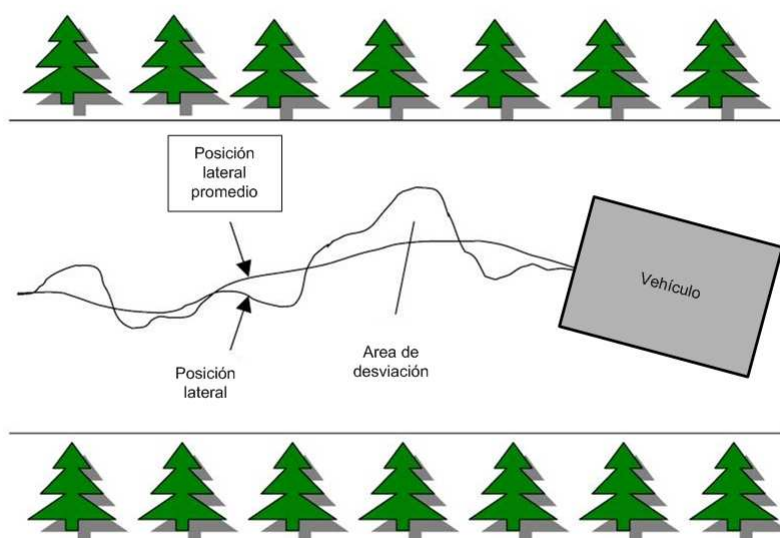


Figura 2.2: Patrón de conducción a partir de la posición del vehículo en la carretera.

2.1.2. Análisis basados en los cambios físicos de los ojos y las expresiones faciales usando procesamiento de imágenes

En el segundo caso se utilizan técnicas de procesamiento de imágenes para detectar los cambios físicos generados sobre los ojos y el rostro del conductor durante el proceso de conducción, ver figura 2.3.

Los métodos basados en el procesamiento de imágenes son altamente fiables con la principal característica de ser no invasivos por lo que no generan molestias ni incomodidad



Figura 2.3: Análisis del rostro con una cámara (Fuente: Heitmann *et al.* [4]).

al conductor; sin embargo, también presentan varios inconvenientes; los mismos que serán analizados más adelante como parte de esta investigación.

Por el momento, se introduce la definición del índice que será utilizado para medir el estado del conductor bajo el análisis de imágenes con técnicas de visión por computador.

2.1.2.1. Perclos

PERCLOS son las siglas de **PER**centage of the Time **Eyelids** are **CLOS**ed. Este índice mide el porcentaje de cierre de los ojos sobre un intervalo de tiempo, excluyendo el tiempo gastado en el cierre normal. En varias investigaciones desarrolladas por la National Highway Traffic Safety Administration (NHTSA) [108], [109] se ha demostrado que es el más adecuado para caracterizar la somnolencia del conductor a través del análisis de los ojos.

2.1.3. Análisis basados en el cambio de las medidas fisiológicas

En el último caso se consideran los métodos enfocados en medir los cambios fisiológicos tales como la variabilidad del ritmo cardíaco, el electromiograma y electroencefalograma del conductor, como un medio para detectar los estados cognitivos humanos. Las ventajas que presentan es la precisión en el diagnóstico del estado. Mientras que como uno de los mayores inconvenientes están en el uso de sensores conectados al cuerpo del conductor [12],[96], [143], [135], [8], [26]. Por ejemplo, el electroencefalograma se obtiene al conectar varios electrodos a la cabeza para registrar las ondas cerebrales, esto se muestra en la figura 2.4 y es el utilizado con mayor frecuencia dentro de este grupo.

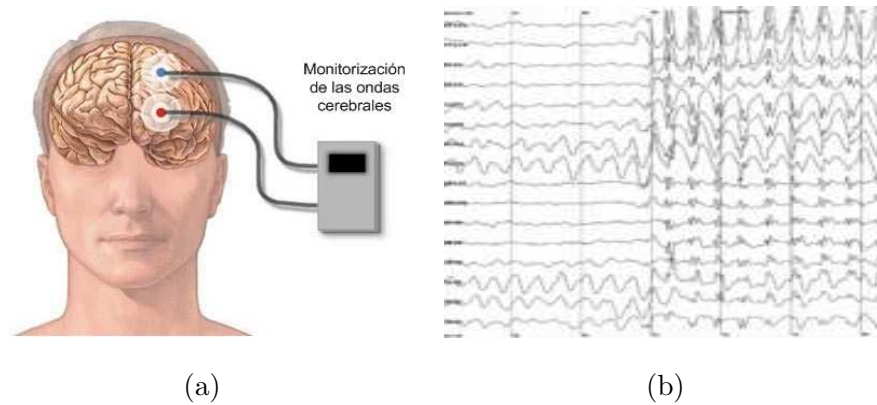


Figura 2.4: Electroencefalograma, (a) aparatos conectados al cerebro del paciente, (b) registro de las ondas cerebrales.

2.2. Métodos basados en el análisis de los ojos y expresiones faciales

De acuerdo a los resultados de varios investigadores, los sistemas de monitorización basados en el análisis de las imágenes de los ojos y las expresiones faciales son difíciles de implementar debido a varios factores complicados de controlar en un vehículo, pero a pesar de ello, constituyen la metodología más apropiada para tratar el problema del análisis del conductor y en particular su somnolencia [96], [130], [12].

El elemento fundamental de este tipo de sistemas lo constituye el sistema de visión embarcado en el propio vehículo, ya que proporciona la información de la cabina donde se aloja el conductor. Otra de las principales propiedades que los caracterizan son la alta fiabilidad, más la característica de ser no invasivos y no intrusivos, además de ser relativamente económicos, tomando en cuenta que, su coste irá decreciendo en la medida que la tecnología se vaya desarrollando.

Los sistemas propuestos en los distintos trabajos que utilizan esta metodología se pueden clasificar de varias maneras; en este caso concreto se lo hará por el tipo de iluminación, porque parece la más adecuada dada la complejidad del problema, las condiciones de conducción (diurna y nocturna) y la bibliografía recopilada durante la investigación. Así se tienen:

-
1. Métodos en el espectro visible
 2. Métodos en el infrarrojo cercano

En el primer caso se utilizan, generalmente, cámaras sencillas junto con varias técnicas clásicas del procesamiento de imágenes (información del color de la piel, correlación, etc.) para detectar la posición de los ojos dentro de la imagen, en conducción diurna. Mientras que en el segundo se utilizan complejos sistemas de visión basados en iluminación infrarroja, para la conducción nocturna.

En base a esta clasificación se desarrollará el estado del arte, donde se pretende dar una visión global del estado de la tecnología y plantear sus virtudes y deficiencias.

2.2.1. Métodos en el espectro visible

Horng *et al.* [153] presentan un sistema que utiliza la información del color de la piel para localizar el rostro, trabajando con el espacio de color HSI , sobre el cual construyen un modelo para identificar el color de la piel humana. Posteriormente, con la información de los bordes localizan los ojos para obtener los modelos a seguir, y utilizando correlación construyen un método para su seguimiento (dynamical template matching, DTM). Además, usando la información del color de los ojos identifican su estado, abierto o cerrado, y calculan el índice de somnolencia. En esta misma tendencia se encuentra el trabajo de Tian y Qin [163] al utilizar la información del color de la piel como parte fundamental de su sistema. En este caso el espacio de color elegido es el $YCbCr$; el mismo que es ampliamente utilizado para identificar zonas de piel, en donde se presta especial atención a las componentes Cb y Cr que son las que mejor capturan las características que representan al color de la piel. Las funciones de proyección vertical y horizontal son empleadas para localizar el rostro y los ojos, respectivamente. Luego, utilizando la función de complejidad (2.1) sobre la imagen del contorno de los ojos el sistema calcula el estado del conductor.

$$com(k) = \sum_{j=1}^m \sum_{i=1}^{n-1} |b(i, j) - b(i-1, j)| * k(i, j) \quad (2.1)$$

donde $b(i, j)$ es el valor de la imagen binaria que contiene el contorno de los ojos en la posición (i, j) , $k(\cdot, \cdot)$ es una función núcleo, m y n corresponden al ancho y alto de la

imagen, respectivamente.

Otro trabajo de relevancia es el de Dong y Wu [152] que también aprovechan las propiedades del espacio de color $YCbCr$; en este trabajo, estos investigadores construyen un modelo de probabilidad del color de piel basado en la función de distribución normal bivalente sobre las componentes Cb y Cr . Una vez localizados los ojos procede a calcular el índice de somnolencia utilizando la distancia entre los párpados para clasificar el estado del ojo entre abierto y cerrado.

Por otra parte, Rongben *et al.* [154] trabajan con el espacio cromático rg sobre el cual construyen un modelo de probabilidad para modelar el color de la piel utilizando la función de distribución gaussiana bivalente. Basados principalmente en el color localizan el rostro y posteriormente los ojos utilizando restricciones sobre el tamaño y la posición. Utilizando una red neuronal artificial (RNA) clasifican el estado del ojo, para ello extraen las características más sobresalientes del ojo con ayuda del wavelet de Gabor. Para el seguimiento de los ojos utilizan el método de correlación, similar al trabajo de Horng *et al.* [153], junto con la distancia de Hausdorff. La distancia de Hausdorff entre dos conjuntos $A = \{a_1, \dots, a_m\}$ y $B = \{b_1, \dots, b_n\}$ se define:

- distancia de A hacia B :

$$d(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (2.2)$$

- distancia de B hacia A :

$$d(B, A) = \max_{b \in B} \min_{a \in A} \|a - b\| \quad (2.3)$$

donde $\|\cdot\|$ es la norma en L^2 .

D’Orazio *et al.* [140] presentan un sistema de vigilancia del conductor que localiza la posición de los ojos para identificar el estado de atención y fatiga en el conductor. El algoritmo propuesto busca sobre toda la imagen regiones que presentan formas elipsoidales (iris del ojo) utilizando para ello la transformada circular de Hough sometida a ciertas restricciones de tamaño.

Jiangwei *et al.* [18], [19] presentan dos trabajos en el campo del análisis del conductor, el primero constituye un método para analizar el estado de los ojos empleando una red de neuronas que toma como entradas varias características geométricas del ojo. Las características se obtienen directamente de la imagen del ojo y se representan a través del siguiente vector:

$$Z_{ojo} = (H_{max}, W_{max}, H_{centro}, h_{L3}, h_{R3}, H_C) \quad (2.4)$$

donde H_{max} es la distancia máxima entre el párpado superior e inferior, W_{max} es el ancho máximo del ojo, H_{centro} es la distancia entre el párpado superior y el centro del iris. Además, el ojo forma un arco cuando se encuentra en estado normal, una línea al estar adormecido y una concavidad en sueño. A fin de expresar estas características se definen tres valores h_{L3} , h_{R3} , H_C . En la figura 2.5 (a) se observa gráficamente el vector descrito. Los estados que se identifican sobre los ojos son: despierto, adormecido y dormido.

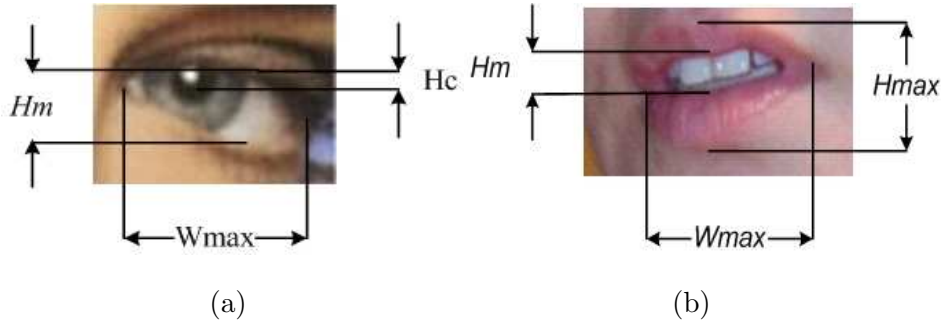


Figura 2.5: Sistema de Jiangwei, (a) parámetros calculados sobre el ojo, (b) parámetros calculados sobre la boca.

En su segunda aportación, Jiangwei *et al.* [19] trabajan sobre el análisis del estado de la boca. Al igual que en el caso anterior, construyen un vector de características con información obtenida de la boca para entrenar una red de neuronas.

$$Z_{boca} = (H_{max}, W_{max}, H_m) \quad (2.5)$$

donde H_{max} corresponde a la distancia máxima entre el labio superior e inferior, W_{max} corresponde al ancho máximo de la boca y H_m corresponde a la distancia entre el labio superior e inferior cuando la boca se encuentra en actividad. Los estados en los que se

2.2 Métodos basados en el análisis de los ojos y expresiones faciales

puede encontrar la boca son tres: dormido, hablando y cerrado. Al igual que en el caso anterior, en la figura 2.5 (b) se aprecia los parámetros descritos.

Para localizar los ojos y la boca, en sus dos aportaciones, Jiangwei *et al.* utilizan el espacio de color dicromático rg y la función de distribución normal bivalente.

Siguiendo con la descripción de los sistemas propuestos se tiene el trabajo de Brandt *et al.* [139] que aprovechan el algoritmo de Viola & Jones [120], [118] para detectar la posición del rostro. Una vez extraída la zona de interés que corresponde al rostro y bajo el supuesto que los ojos corresponden a la zona más oscura obtienen una imagen binaria donde se localizan los ojos. Posteriormente, utilizan el algoritmo de flujo óptico (optical flow) sobre la zona de los ojos para estimar su movimiento y el de la cabeza.

Branzan *et al.* [1] también presentan un sistema de detección de la somnolencia que utiliza una cámara de bajo coste para capturar las imágenes del conductor. Para conocer el estado de los ojos utilizan el método de correlación (template matching) en conjunto con dos modelos (ojos abiertos y ojos cerrados) obtenidos durante el proceso de calibración. En la figura 2.6 se presenta el sistema descrito.



Figura 2.6: Sistema presentado por Branzan *et al.* para la detección de la somnolencia.

Hayami *et al.* [142] construyen un sistema para detectar somnolencia a través del análisis de los ojos, introduciendo un nuevo índice que se denomina *tasa de movimientos verticales del ojo*, que depende del índice Perclos y que viene dada por la siguiente ecuación:

$$V = \frac{V_c}{1 - P} \quad (2.6)$$

donde V_c es el número de movimientos verticales del ojo y P es el valor de Perclos [108]. V denota la frecuencia del movimiento vertical del ojo para 60, 30 y 15 segundos cuando el ojo esta abierto. Este trabajo lo desarrollan sobre un simulador que consta de un casco y un soporte que alberga una cámara que adquiere directamente las imágenes de la zona de los ojos, con lo cual, no es necesario ningún método de detección; dicho sistema es presentado en la figura 2.7.

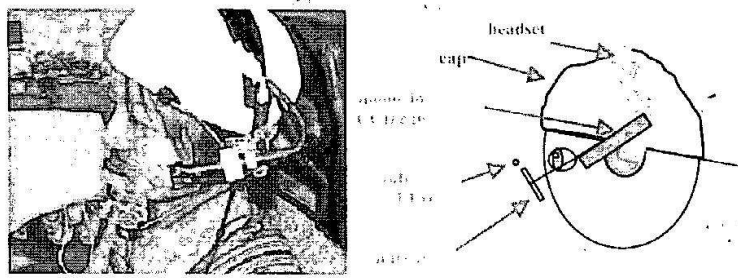


Figura 2.7: Sistema desarrollado por Hayami *et al.* para capturar las imágenes de los ojos.

2.2.2. Métodos en el infrarrojo cercano

Entre las principales investigaciones que se pueden citar en este caso se encuentran los trabajos desarrollados por Ji y Yang [123], [124], Ji *et al.* [125], Zhu *et al.* [164], [165] donde se presenta un sistema de detección de la somnolencia basado en iluminación infrarroja y visión estéreo. Para localizar la posición de los ojos se utilizan las propiedades físicas que presenta la pupila frente a la iluminación infrarroja. En el rango comprendido entre los 700 nm. y 900 nm. la pupila adquiere un brillo característico que lo diferencia del resto de objetos y que permite reducir el trabajo para su localización. El sistema de iluminación infrarrojo encargado de filtrar la luz y resaltar la pupila se presenta en la figura 2.8.

El efecto de este proceso físico se aprecia en la figura 2.9; es así que en la figura 2.9 (a) se presenta la imagen donde se resalta la pupila; mientras en la figura 2.9 (b) se observa la imagen donde se opaca la pupila.

2.2 Métodos basados en el análisis de los ojos y expresiones faciales

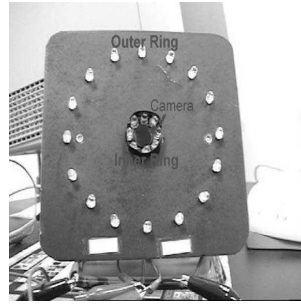


Figura 2.8: Sistema de iluminación diseñado por Ji *et al.* para desarrollar sus trabajos.

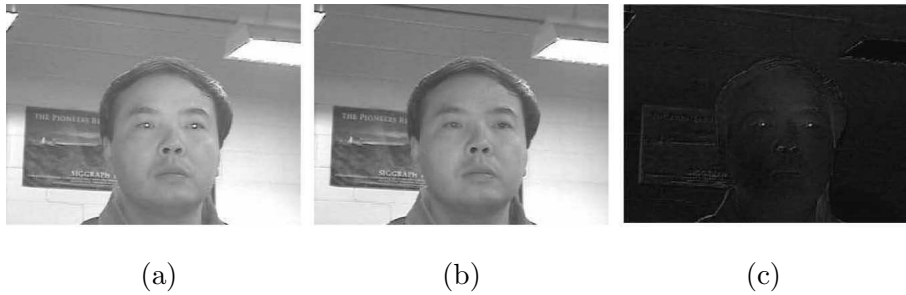


Figura 2.9: Diferencia de imágenes del campo par e impar utilizado por Ji *et al.*

La imagen de la figura 2.9(c) es la que recibe atención especial, porque sobre ella se resaltan las zonas de mayor probabilidad donde se pueden localizar los ojos. Esta última imagen se obtiene por el valor absoluto de la diferencia de las imágenes (a) y (b) de la figura 2.9.

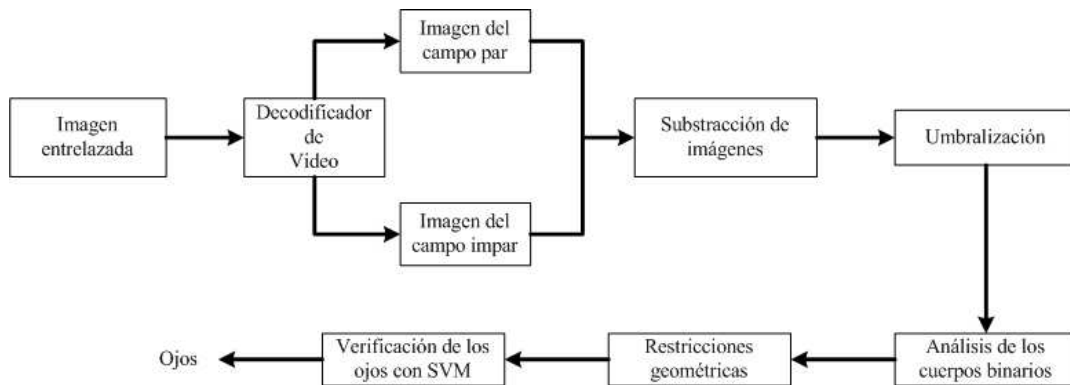


Figura 2.10: Diagrama para la detección de los ojos utilizado por Ji *et al.*

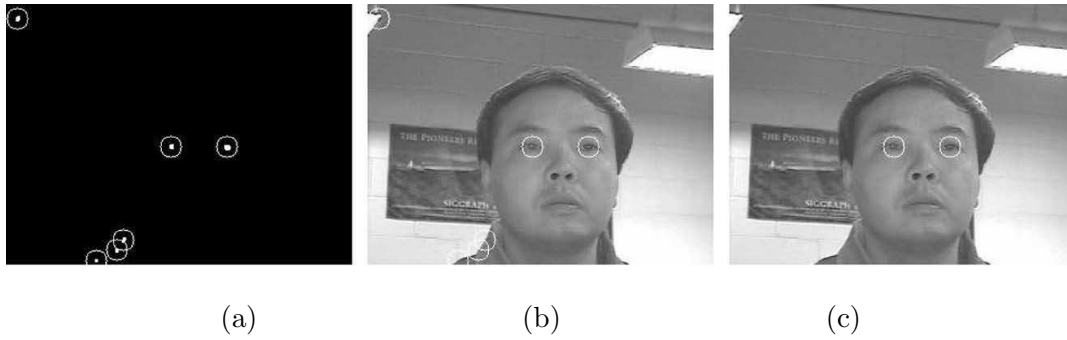


Figura 2.11: Resultados de detección de los ojos utilizado por Ji *et al.*

Tras varios pasos en los que se incluye un exhaustivo proceso de búsqueda y clasificación se localizan los ojos. La figura 2.10 muestra el esquema para localizar su posición, que una vez detectados se presentan en la figura 2.11 (c), donde se indica el resultado final.

Bergasa *et al.* [78], [79] han desarrollado un sistema no invasivo que también utiliza iluminación infrarroja, su sistema calcula el nivel de vigilancia del conductor usando un autómata de estado finito [116] con seis estados. Los estados que maneja el autómata son: (i) estado inicial, (ii) estado de seguimiento, (iii) estado de pérdida de seguimiento, (iv) estado abierto, (v) estado cerrado, y (vi) estado cerrando. El estado *cerrando* equivale a decir semi-abierto o semi-cerrado. Con la ayuda del autómata construyen varios índices, entre ellos, Perclos. Posteriormente, para ampliar las prestaciones que ofrece este sistema, sus autores desarrollan un método para detectar la distracción del conductor, a través del análisis de la posición del rostro.

Otro trabajo importante y ampliamente citado en la literatura corresponde al desarrollado por Grace *et al.* [129], [40] quienes han construido un sistema para ser instalado en vehículos de carga pesada (trailers y camiones). Grace también utiliza iluminación infrarroja para localizar la posición de los ojos, su sistema captura a través de dos cámaras dos imágenes, cada cámara tiene incorporado un filtro, una cámara compuesta por un filtro de 850 *nm* y la otra por el filtro de 950 *nm*. A 850 *nm* la retina refleja el 90 % de la luz infrarroja incidente sobre la retina y a 950 *nm* refleja únicamente el 40 % de la luz infrarroja incidente sobre la retina. El efecto es el mismo que se presentó en la figura 2.11 de Ji *et al.* Una vez localizados los ojos, el sistema se encarga de calcular el índice Perclos

2.2 Métodos basados en el análisis de los ojos y expresiones faciales

para estimar el grado de somnolencia del conductor.

Por su parte Liu *et al.* [155] presentan una solución basada en la media y la varianza de la imagen obtenida con iluminación infrarroja para calcular un umbral que separa los potenciales ojos del resto de la imagen. El umbral que proponen viene dado por la ecuación (2.7), que es fácil de implementar y que requiere un coste computacional mínimo.

$$T = 255 - \mu - \gamma\sigma \quad (2.7)$$

donde μ y σ son la media y la varianza de la imagen, respectivamente, γ es un coeficiente dado por (2.8).

$$\gamma = 0,9 + \text{ceil}\left(\frac{\max(95 - \mu, 0)}{2}\right) \quad (2.8)$$

donde $\text{ceil}(x)$ retorna el menor entero más grande que x , es decir, redondea al entero más inmediato. Posteriormente para localizar los ojos utiliza un método de clasificación y para el seguimiento utiliza el algoritmo Mean Shift, que trabaja con el histograma de la zona de interés extraída de la imagen [53].

A pesar de las grandes ventajas que presenta la iluminación infrarroja para resaltar las pupilas y con ello detectar la posición de los ojos, también presenta ciertas desventajas que hace que no sea suficiente para trabajar en ambientes sometidos a los cambios de iluminación provenientes de las condiciones ambientales. El brillo de las pupilas desaparece cuando la iluminación externa interfiere sobre el rostro [140], también lo hace cuando el conductor gira su rostro o cierra sus ojos [165]. Es decir, la iluminación infrarroja funciona correctamente bajo condiciones estables de iluminación [155] y cuando el rostro se encuentra perpendicular al rayo de luz que emite el sistema de iluminación. En resumen, el brillo y el tamaño de las pupilas es función de la orientación del rostro, de las condiciones externas de iluminación y de la distancia del conductor a la cámara. Esto tiene como consecuencia que el sistema sea altamente inestable por depender de factores que no se pueden controlar. Sin embargo, para corregir los problemas de iluminación y mantener el seguimiento de los ojos, Zhu *et al.* [165], Ji *et al.* [125] y Liu *et al.* [155] han utilizado una doble estrategia que consiste en utilizar el filtro de Kalman y el algoritmo Mean Shift, con

lo que han logrado mejoras en sus sistemas. Cabe mencionar que este último algoritmo debido a que utiliza la información del histograma [53] es altamente dependiente de los cambios de iluminación, con las consecuencias que ello conlleva.

2.3. Sistemas comerciales

Las principales empresas vinculadas al sector automotriz están trabajando o participando, directamente o a través de sus empresas fabricantes de componentes, en la construcción de sistemas de ayuda a la conducción. Entre los sistemas que destacar se encuentran los sistemas de vigilancia del conductor; es así, que en las siguientes secciones se describirán varias empresas que utilizan las distintas estrategias planteadas:

- Técnicas del análisis de los ojos y expresiones faciales
- Análisis de la carretera
- Información de las señales fisiológicas

2.3.1. Attention Technologies

La empresa Attention Technologies Inc. [68] presenta y comercializa un sistema para la detección de la somnolencia. En la actualidad se registran dos versiones: Copilot, figura 2.12, y Driver Fatigue Monitoring DD850, figura 2.13.

Copilot y DD885 han sido desarrollados y construidos por Grace *et al.* [129], bajo un proyecto conjunto desarrollado por la Universidad Carnegie Mellow y la NHTSA [109]. Los dos sistemas están diseñados para trabajar en vehículos de carga pesada (camiones) y calculan el índice de somnolencia Perclos, en tiempo real. La base fundamental de los dos sistemas lo constituye el sistema de visión con iluminación infrarroja.

2.3 Sistemas comerciales

- **Copilot:**



Figura 2.12: Copilot: (a) vista general del sistema, (b) entorno virtual de implementación sobre un camión.

- **Driver Fatigue Monitoring DD850:**



Figura 2.13: Driver Fatigue Monitoring: DD850.

2.3.2. Seeing Machine

Seeing Machine [102] es una empresa dedicada a desarrollar y proveer productos de alta tecnología para el análisis y seguimiento del rostro y los ojos utilizando técnicas basadas en Visión por Computador.



Figura 2.14: Sistema propuesto por Seeing Machine: (a) sistema de visión, (b) simulador virtual para analizar el comportamiento del conductor.

Como parte del proyecto TREV (Transport Research Experimental Vehicle) [80], su principal producto es el denominado *faceLab* que se muestra en la figura 2.14. El sistema consiste de un sistema de visión estéreo que utiliza iluminación infrarroja.

faceLab es un sistema que está siendo utilizado en varias aplicaciones, entre ellas: biometría, estudios de psicología, y por supuesto, sistemas de vigilancia del conductor, por ejemplo, Kutila [96] utilizó este sistema para elaborar su tesis doctoral, donde construye patrones de conducción.

2.3.3. Grupo Bosh

La empresa Bosh [59] se encuentra trabajando en el diseño y construcción de sistemas de interface humano-máquina (HMI); entre los cuales, se encuentra el análisis del conductor. Su sistema de detección de la somnolencia consiste en un simulador construido para desarrollar varios experimentos. En la figura 2.15 se observa el simulador propuesto por ésta empresa.

2.3.4. Volvo Cars

Recientemente, la empresa automotriz Volvo ha empezado a incluir en sus versiones de lujo un sistema de seguridad preventiva que da apoyo al conductor y lo alerta cuando comienza a presentar síntomas de cansancio o desconcentración [23].

2.3 Sistemas comerciales

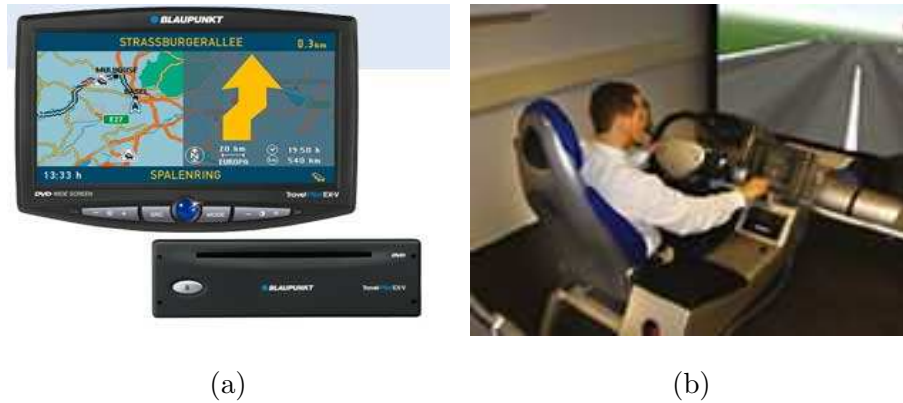


Figura 2.15: Simulador utilizado por Bosh para estudiar al conductor.

Mediante una cámara y un sensor de radar instalados en la parte frontal del vehículo, extrae información para calcular la dirección de la carretera, figura 2.16 (a), y junto con los datos del volante genera un patrón de conducción que es comparado con uno en conducción normal. El sistema le avisa al conductor cuando éste pierde el control mediante una señal acústica, al mismo tiempo, en el panel de instrumentos aparece un mensaje de alerta, figura 2.16 (b).



Figura 2.16: Control de alerta para el análisis del conductor.

2.3.5. Mercedes-Benz Anti-Fatigue Driver Alarm

El sistema experimental de Mercedes-Benz recurre a un sistema de visión infrarrojo que monitoriza los ojos del conductor, evaluando su nivel de somnolencia. Además, a través de un sistema de sensores colocados en la cabeza, evalúa la actividad cerebral del conductor



Figura 2.17: Sistema experimental para la detección de la somnolencia de Mercedes-Benz.

mediante un encefalograma, de manera que combinando toda ésta información de manera adecuada, predice el estado de somnolencia del conductor.

En la figura 2.17 se observa una panorámica del sistema que pretende comercializar en el futuro cercano, como parte de un sistema de prevención de accidentes.

2.3.6. AWAKE

En este proyecto intervienen empresas privadas y algunos sectores públicos europeos con la intención de construir un sistema de vigilancia del conductor. Entre las empresas privadas se encuentran Siemens VDO, Navigation Technologies, DaimlerChrysler, etc.; y entre las empresas públicas presentes se encuentran la Swedish National Road and Transport Research Institute, Belgian Road Safety Institute, la Universidad de Stuttgart, etc. En la figura 2.18 se aprecia la estructura del sistema desarrollado por este consorcio.

Los objetivos que han motivado el desarrollo de este SAAC es el contribuir en la redu-

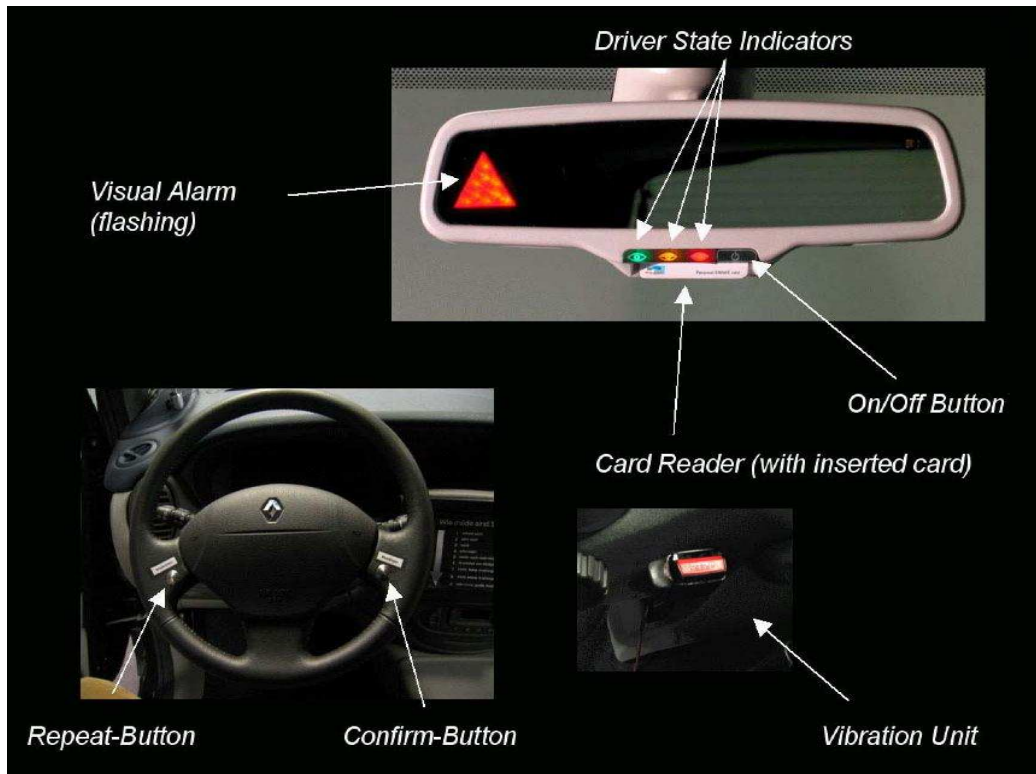


Figura 2.18: Esquema del sistema desarrollado en el proyecto AWAKE (2.002).

cición del número de accidentes de tráfico relacionados con la somnolencia, aumentando la seguridad en la carretera. Su módulo principal analiza los ojos y calcula varios parámetros para detectar la presencia de somnolencia, en este sentido, incorpora información del radar, del GPS, etc., para mejorar las estimaciones del estado de conductor.

2.3.7. Mitsubishi Motors Corporation

Esta empresa presenta un vehículo que sirve como plataforma experimental para probar los SAAC que se encuentra desarrollando [104].

Uno de sus últimos modelos consta de un sistema de detección del cambio involuntario de carril (Lane Keep Assist System) formado por una cámara que analiza las líneas de la carretera y a través de ello estima el estado del conductor. En la figura 2.19 se observa el vehículo que utilizan para realizar sus experimentos.



Figura 2.19: Vehículo de Mitsubishi para el desarrollo de SAAC.

2.3.8. AssistWare Technology

La empresa AssistWare Technology [145] se dedica al desarrollo de tecnología para ayudar en la reducción de los accidentes de tráfico, es así que en 1.996 presentó el sistema SafeTRAC, el mismo que se muestra en la figura 2.20 (a).

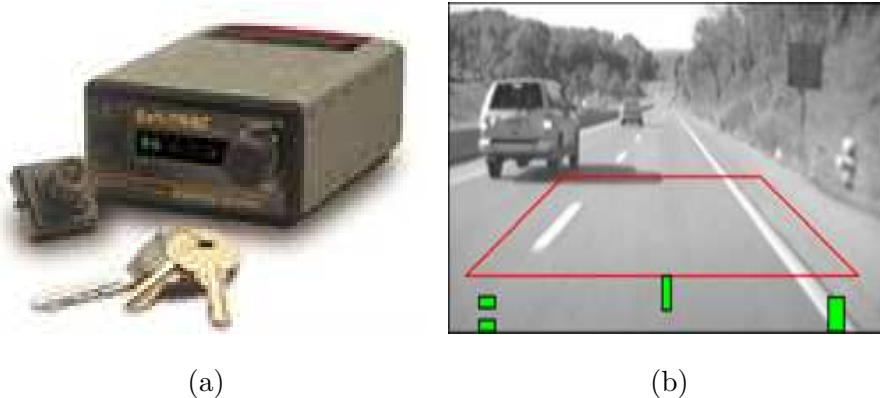


Figura 2.20: Sistema SafeTrac 125, (a) sistema de visión, (b) imagen de la carretera.

El sistema analiza la carretera utilizando la información visual obtenida a través de una cámara. Con la información de la zona de la carretera que corresponde a aproximadamente 50 metros al frente del vehículo, figura 2.20 (b), el sistema calcula la posición y la trayectoria para estimar dos situaciones de peligro: somnolencia y salida involuntaria del carril. A través de una alarma sonora indica si el conductor se encuentra en alguna de las situaciones de peligro.

2.3 Sistemas comerciales

Desafortunadamente, este sistema no es adecuado para vías en mal estado, donde las líneas de la carretera no son visibles; por lo que sería difícil estimar un modelo como el propuesto.

2.3.9. Delta Tooling

El último accesorio de seguridad activa para el automóvil es un asiento que detecta el adormecimiento del conductor. Este dispositivo está siendo desarrollado por la Universidad de Tokio en colaboración con la Universidad de Oita, el Instituto Shimane de Ciencias de la Salud y el fabricante Delta Tooling.

Este diseño utiliza varios sensores instalados en el asiento, para detectar los signos fisiológicos que señalan la aparición del sueño, en este sentido, el sistema se encuentra formado por dos sensores que controlan el pulso y dos que miden la respiración.

Sus desarrolladores proyectan tenerlo listo dentro de cinco años, tiempo en el cual, planean tener una versión con alarma despertadora que responda automáticamente a los síntomas del sueño. En la figura 2.21 se observa el prototipo de este sistema.



Figura 2.21: Asiento que detecta el sueño.

2.4. Limitaciones de los sistemas comerciales actuales

Se han presentado varios prototipos desarrollados por empresas especializadas en la construcción de sistemas de vigilancia del conductor, sus diseños en la mayoría de casos se encuentran en etapa de desarrollo y pruebas, con resultados generados en laboratorios y/o en entornos de realidad virtual. En consecuencia, muchas de estas empresas esperan tener sus diseños operativos en un par de lustros o décadas, y en el mejor de los casos, presentan versiones de prueba que prestan servicios limitados y que tan solo se encuentran instalados en sus modelos de lujo.

2.5. Conclusiones

En este capítulo se han descrito las principales técnicas utilizadas para analizar la somnolencia del conductor, se las ha dividido en tres grupos: métodos basados en patrones de conducción, análisis de los ojos y expresiones faciales utilizando Visión por Computador y análisis de las señales fisiológicas. Cada una de ellas abarcan distintas metodologías para abordar el problema en cuestión y han servido de base para el desarrollo de distintos sistemas.

Como es de especial interés, se ha puesto énfasis en las técnicas basadas en el análisis de los ojos y el procesamiento de imágenes. Al dividir las por el tipo de iluminación se ha conseguido segmentarlas en dos grupos claramente definidos, métodos en el espectro visible (Horng *et al.* [153], Tian *et al.* [163], Dong and Wu [152], Rongben *et al.* [154], D'Orazio *et al.* [140], Jiangwei *et al.* [18], [19], Brandt *et al.* [139], Branzan *et al.* [1], Hayami *et al.* [142]) y métodos en el infrarrojo cercano (Ji *et al.* [123], [124], [125], Zhu *et al.* [164], [165], Bergasa *et al.* [78], Grace *et al.* [129], Ayoob *et al.* [40]). Posteriormente se ha complementado la información describiendo las técnicas específicas que cada investigador ha utilizado para dar solución al problema de la somnolencia.

La mayoría de las investigaciones expuestas utilizan el índice Perclos para detectar la somnolencia del conductor. Perclos ha demostrado su robustez y su versatilidad y ha sido el único índice reconocido y probado como efectivo por la NHTSA de Estados Unidos [108].

2.5 Conclusiones

En esta misma tendencia, Hayami *et al.* [142] presentan un nuevo índice, denominada “tasa de movimientos verticales del ojo”, que en definitiva es una variante de Perclos. Por su parte, Horng *et. al.* [153] y Dong and Wu [152] proponen un umbral de cinco fotogramas consecutivos para detectar la somnolencia.

En la misma línea de investigación las empresas privadas están trabajando activamente en el desarrollo de distintos sistemas de detección de la somnolencia del conductor. En este sentido, las industrias vinculadas al sector automotriz se encuentran dirigiendo sus intereses hacia los sistemas de ayuda a la conducción, particularmente, a la implementación de sistemas basados en el procesamiento de imágenes y complementados en muchos casos con otras técnicas, como el análisis de las señales fisiológicas.

Lo anterior demuestra que el tema de la seguridad activa o preventiva va adquiriendo importancia y de allí los grandes esfuerzos dedicados para desarrollar sistemas avanzados de ayuda a la conducción que vigilen y analicen el comportamiento del conductor (directa e indirectamente), permitiendo que la tecnología contribuya en la reducción de los accidentes de tráfico producto de los errores humanos, sean estos voluntarios y/o involuntarios.

Finalmente, se puede hacer referencia a que los SAAC aún se encuentran en etapa de investigación, desarrollo y pruebas, por lo que pasará al menos un par de décadas hasta que lleguen a ser incorporados en los vehículos de venta cotidiana, donde existe un mercado potencial al cual se puede acceder incorporando seguridad y confort adicionales para mejorar las condiciones de conducción y, a la vez, generando información oportuna para que el conductor tome las decisiones acertadas con suficiente antelación.

Apreciación Global y Sistema de Percepción

En este capítulo se muestra la apreciación global de lo que constituye la tesis, a partir de ello, se exhiben los escenarios para elaborar el software y el hardware que componen el sistema de detección de la somnolencia desarrollado en la Universidad Carlos III de Madrid. Con estricto rigor, se presenta la información de los sistemas de visión utilizados en el día y la noche, relacionados con la conducción diurna y nocturna. Adicionalmente, se bosqueja la estructura de los algoritmos para los dos escenarios.

3.1. Apreciación global y características generales

En el capítulo uno, sección 1.4, se indicó que existen dos intervalos de tiempo donde ocurren con mayor frecuencia los accidentes relacionados con la somnolencia, el primero se sitúa en las primeras horas de la tarde y el segundo comprende las horas de la madrugada.

Esta información indica que el sistema que se proyecta construir debe funcionar en los dos escenarios, es decir, el día y la noche, relacionados directamente con la conducción diurna y nocturna, respectivamente; exigiendo que sea necesario utilizar sistemas de visión

y construir algoritmos que funcionen en las dos situaciones.

En lo referente al software, de manera global el sistema que se presenta en esta tesis esta formado de los siguientes módulos:

- **Detector del rostro:** encargado de detectar la posición del rostro del conductor dentro de la imagen
- **Detector de los ojos:** localiza la posición de los ojos en la imagen
- **Detector del estado de los ojos:** analiza el estado de los ojos y generar el índice de somnolencia
- **Detector de la distracción:** estima si el conductor se encuentra distraído mientras conduce
- **Módulo de seguimiento:** diseñado para seguir los ojos y el rostro, permitiendo mantenerlos localizados, reducir el espacio de búsqueda y por ende el tiempo de cómputo

La figura 3.1 indica globalmente la interacción entre los módulos, este diagrama es válido para los dos sistemas (día y noche), a pesar que sufrirá modificaciones en función de las necesidades específicas de cada escenario.

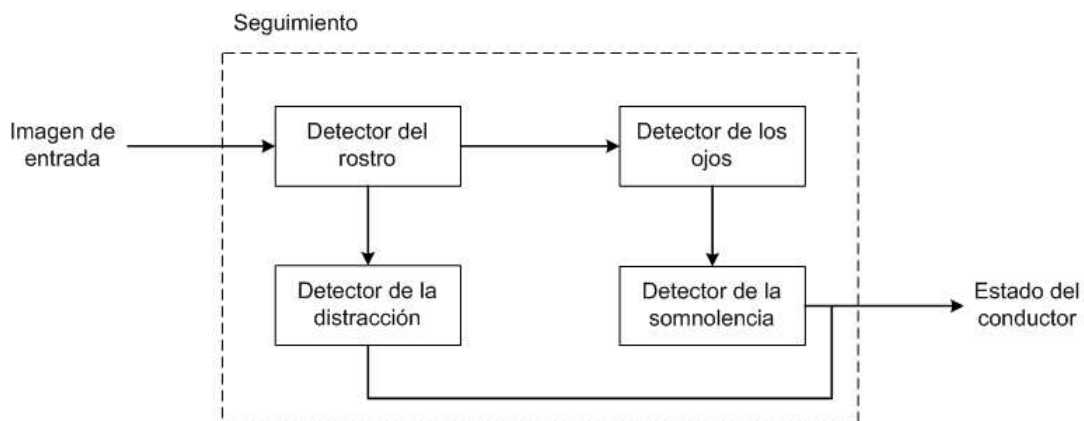


Figura 3.1: Apreciación global del sistema de detección de la somnolencia.

3.1 Apreciación global y características generales

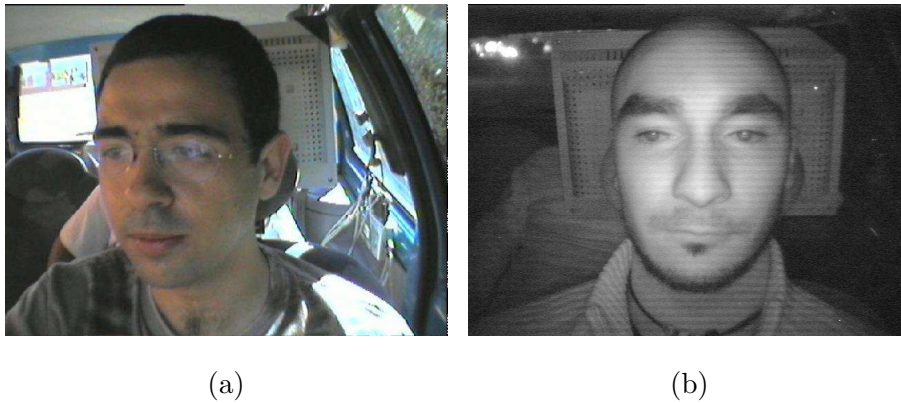


Figura 3.2: Imágenes capturadas con el sistema de visión del vehículo Ivvi, (a) imagen en el espectro visible, (b) imagen en el infrarrojo cercano.

En el día, primero se detecta el rostro, se continúa con la detección de los ojos, el seguimiento, el cálculo del índice de somnolencia y se finaliza con el análisis de la distracción. Mientras que en la noche, se comienza por detectar los ojos, se continúa con la detección del rostro, el seguimiento, el cálculo del índice de somnolencia y se finaliza con el análisis de la distracción.

Siguiendo con la descripción, la figura 3.2 muestra dos imágenes tomadas con el sistema de percepción de la plataforma Ivvi, en ellas se puede apreciar la diferencia y similitud existente en los escenarios día y noche, lo que influirá y guiará el diseño de los algoritmos.

Por lo anterior, el sistema de detección de la somnolencia de la Universidad Carlos III de Madrid se compone de dos partes. La primera corresponde a lo que se denominará el *sistema de detección de la somnolencia en el espectro visible*, y la segunda será el *sistema de detección de la somnolencia en el infrarrojo cercano*.

Para trabajar en el espectro visible se utiliza un sistema de visión compuesto principalmente por una cámara, mientras que para el caso restante se construye un complejo sistema de visión basado en la iluminación infrarroja. Estos sistemas serán descritos en las próximas secciones.

3.2. Sistema de percepción del vehículo IvvI

En estrecha relación con la división realizada se diseña el hardware que forma parte de este estudio, es así que el sistema de percepción que forma parte de esta tesis se encuentra compuesto por las siguientes partes:

- Sistema de visión en el espectro visible
- Sistema de visión en el infrarrojo cercano

Cada uno de ellos serán analizados en las próximas secciones, donde se presentará en detalle la información que ha contribuido a su construcción.

3.2.1. Sistema de visión en el espectro visible

El sistema de visión artificial empleado en este caso se encuentra formado principalmente por una cámara, conectada a un ordenador, que incluye una tarjeta de adquisición y procesamiento de imágenes. En la figura 3.2 (a) se presentó un ejemplo generado con este sistema.

Finalmente, éste sistema forma parte de un sistema más complejo, que es el sistema de visión con iluminación infrarroja, el mismo que será descrito en la próxima sección.

3.2.2. Sistema de visión en el infrarrojo cercano

Los ojos presentan características visibles que permiten identificar el estado del conductor, de aquí la necesidad e importancia de detectarlo. Aprovechando las propiedades espectrales de la pupila bajo la influencia de la iluminación infrarroja se procede a construir el sistema de iluminación que será el encargado de excitarlas para facilitar su detección, especialmente en la noche.

El sistema se encuentra formado por cuatro módulos que dan origen al sistema de visión bosquejado en la figura 3.3, cada uno de ellos contiene información importante para llegar al diseño final y ellos son:

- Sistema de visión

3.2 Sistema de percepción del vehículo Ivvi

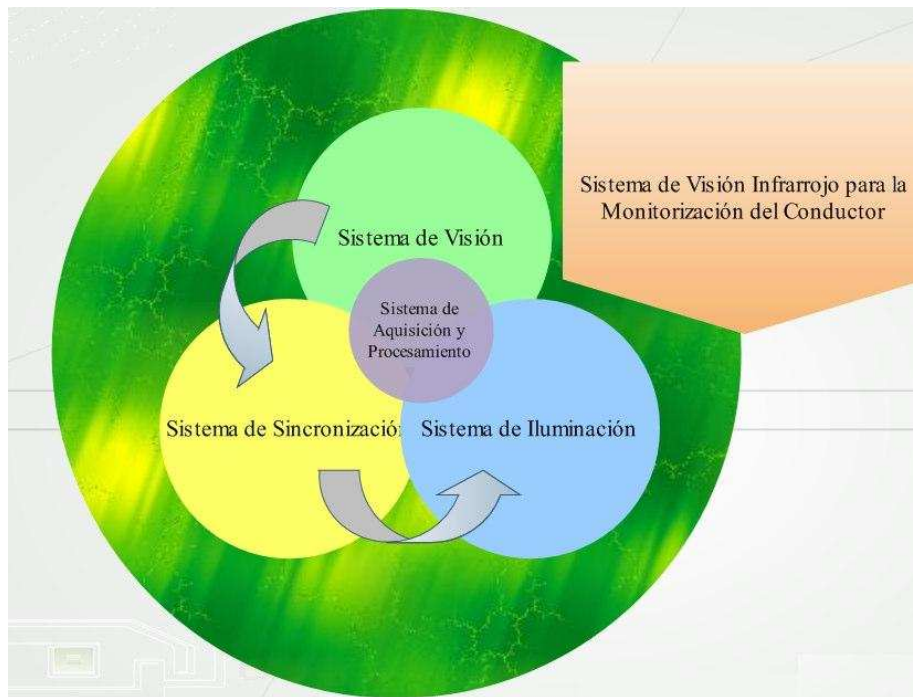


Figura 3.3: Módulos del sistema de visión con iluminación infrarroja.

- Sistema de iluminación
- Sistema de sincronización
- Sistema de adquisición de imágenes

De los sistemas arriba mencionados, los tres primeros son diseñados y construidos en ésta parte de la tesis, mientras que el cuarto se encuentra compuesto exclusivamente por lo que constituye el sistema de visión en el espectro visible.

Los detalles en extenso de la construcción de cada uno de los subsistemas, sus elementos y electrónica asociada se presentan en el apéndice. En esta parte solo se presenta su diseño y resultado final.

3.2.2.1. Espectro electromagnético y la luz infrarroja

El espectro electromagnético es un conjunto de ondas que abarca un amplio abanico de longitudes; que incluyen tanto la luz ultravioleta (cortas longitudes de onda) hasta la luz infrarroja (largas longitudes de onda), pasando por la luz visible (longitudes de

onda intermedias). A lo largo de este amplio rango de frecuencias se manifiestan todos los colores. La figura 3.4 muestra gráficamente este fenómeno, donde se identifica que la luz visible para el ojo humano se encuentra bastante acotada dentro de este espectro.

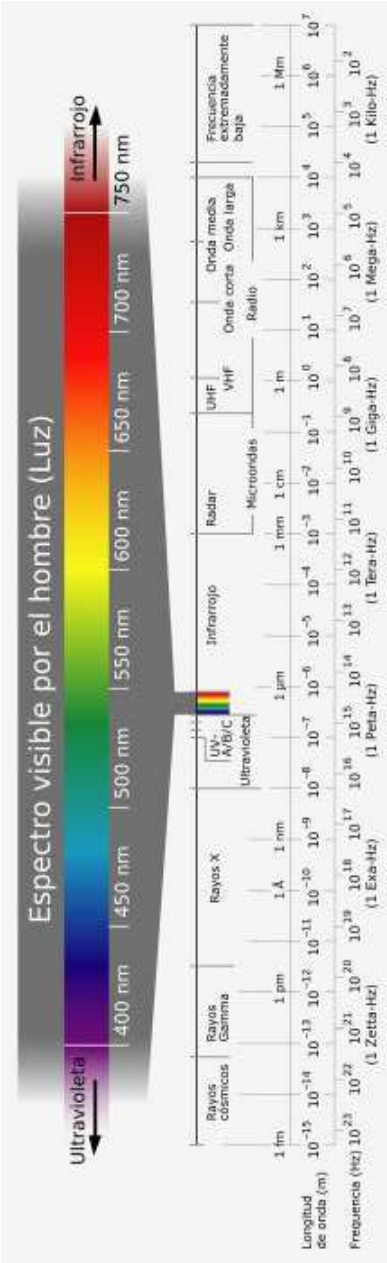


Figura 3.4: Espectro electromagnético.

Los *rayos gamma* y los *rayos X* se localizan al extremo izquierdo del espectro y son los que contienen la mayor energía y a la vez la menor longitud de onda, siguiendo hacia

3.2 Sistema de percepción del vehículo Ivvi

la derecha el recorrido del espectro se encuentra la luz ultravioleta (UV) que comprende las longitudes de onda menores a 400 nm. y mayores a 10nm., aproximadamente. La luz visible se encuentra en el intervalo de los 400 nm. a los 750 nm., también aproximadamente; posteriormente *a partir de los 750 nm. en adelante se encuentran la luz infrarroja* y las ondas de radio, cuyas longitudes de onda son cada vez crecientes y que pasan por las microondas (30 cm.) hasta llegar, por ejemplo, a las ondas de radio de baja frecuencia que aproximadamente se localizan en los 10 km.

3.2.2.2. División de la luz infrarroja

Se podría decir que la iluminación infrarroja es la que ocupa el lugar que se encuentra entre la luz visible y las microondas dentro del espectro electromagnético. Dentro se tiene otra subclasificación, es decir, el infrarrojo se divide en tres regiones espectrales que aproximadamente se encuentran en los siguientes intervalos:

- **Infrarrojo cercano:** entre 700 y 1.500 nm. La luz en esta franja no es térmica, es decir, no está producida por la radiación del calor
- **Infrarrojo medio:** desde 1.500 hasta 15.000 nm.
- **Infrarrojo lejano:** desde los 15.000 nm. hasta 1 mm.

de esta clasificación, el sistema de visión con iluminación infrarroja que se presenta utiliza la iluminación del primer grupo, las razones se expondrán a continuación.

3.2.2.3. Efecto de la luz infrarroja sobre el ojo humano

Las longitudes de onda que mayor peligro conllevan para el ser humano son las de menor longitud de onda; mientras que las de mayor longitud de onda representan pocas dificultades. Según varios estudios una emisión prolongada de cierta potencia en el espectro infrarrojo puede afectar a la retina y a la cornea, principalmente; produciendo photoretinitis e incluso ceguera, ya que pueden recalentar el tejido celular del ojo. En este punto es donde el diseño debe tomar en cuenta estas limitaciones ya que al estar radiando continuamente sobre el ojo se podría producir lesiones ante una potencia descontrolada.

Experimentalmente se ha encontrado que para un espectro comprendido entre los 770 nm. y los 950 nm. el límite de radiación que produce lesiones en la retina (sobre el ojo de un primate) es aproximadamente de $30\text{W}/\text{cm}^2$ para 1 segundo, $23\text{W}/\text{cm}^2$ para 10 segundos, $20\text{W}/\text{cm}^2$ para 100 segundos y $19\text{W}/\text{cm}^2$ para 1.000 segundos.

El American Conference of Governmental Industrial Hygienists (ACGIH) anuncia en The Threshold Limit Values (TLVs) and Biological Exposure Indices (BEIs) que *el límite de irradianza que se puede ofrecer sin causar daño a la cornea y cristalino en una longitud de onda de 700 nm. a 3.000 nm. en emisión led no debe exceder los $10\text{ mW}/\text{cm}^2$ en exposiciones mayores a 1.000 segundos.*

Por lo tanto queda clara la necesidad de utilizar un elemento radiante de iluminación infrarroja en el infrarrojo cercano y que se encuentre limitado en la cantidad de energía radiada.

3.2.2.4. Sistema de visión

El sistema de visión esta compuesto básicamente por una cámara de video y su soporte. La cámara proporciona una imagen analógica en video compuesto, de esta señal se obtiene la señal de sincronismo que será usada en el sistema de sincronización.

3.2.2.5. Sistema de iluminación

El objetivo que se persigue es obtener una imagen donde las líneas del campo par se encuentren iluminadas con una intensidad de luz infrarroja y las del campo impar con otra intensidad de luz. Este efecto se puede conseguir incidiendo con dos sistemas de luz con diferentes ángulos. De esta forma se consigue que las pupilas de los ojos tengan un nivel de luz diferente. Una vez obtenido este efecto y mediante algún procesado se puede ubicar y localizar las pupilas en la imagen.

Como se ha venido indicando, la potencia del sistema de iluminación ha de ser tal que no distraiga al conductor ni tampoco pueda dañar la integridad física de sus ojos.

El sistema de iluminación está situado en el mismo soporte en el que se encuentra la cámara, ver figura 3.5, y consta de un anillo central de leds alrededor de la óptica de la

3.2 Sistema de percepción del vehículo IvvI

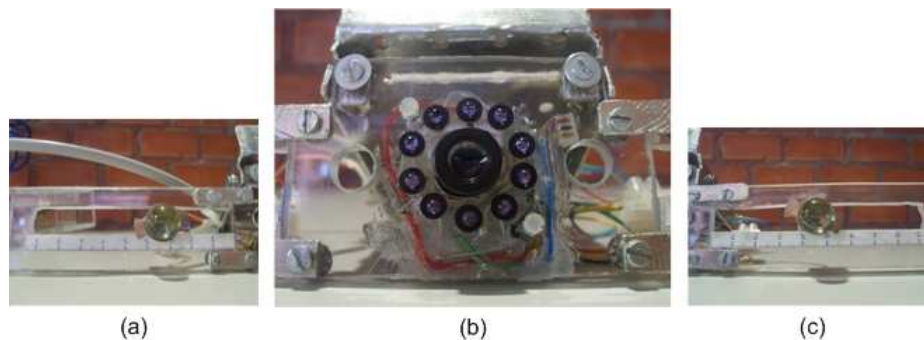


Figura 3.5: Sistema de visión con iluminación infrarroja, (a) anillo de iluminación central, (b-c) iluminación lateral.

cámara y de un grupo de leds en su horizontal, a cada lado de la cámara sobre los soportes auxiliares añadidos.

Este sistema iluminará de forma intermitente al conductor logrando obtener el efecto deseado para el posterior procesado de la imagen. El sistema que regula el momento de disparo es el de sincronización.

3.2.2.6. Sistema de sincronización

Este sistema se encarga de sincronizar el momento en el que el sistema de visión capta las líneas del campo par de la imagen con el momento de disparo del anillo central de iluminación, mientras que las líneas del campo impar serán las que disparen la iluminación proporcionada por los leds exteriores. El eje central de funcionamiento rota sobre el integrado LM1881N que no es mas que un separador de líneas de sincronismo de una señal video compuesto en sistema PAL, NTSC o SECAM. Este integrado toma la señal que proporciona el sistema de visión y la descompone en dos señales que entran en el módulo que controla el disparo del sistema de iluminación. El circuito de sincronismo, junto con el circuito de control de potencia del sistema de iluminación, se ha integrado dentro de un mismo módulo de fácil ubicación, ver figura 3.5.

3.2.2.7. Sistema de adquisición de imágenes

Corresponde plenamente al sistema de visión en el espectro visible.

3.2.2.8. Resultados

El diseño del circuito electrónico se presenta en la figura 3.6, que alberga a los cuatro sistemas y cuyo resultado final es el sistema de visión con iluminación infrarroja presentado en la figura 3.7 e instalado sobre el vehículo IvvI, figura 3.8, generando las imágenes de la figura 3.9.

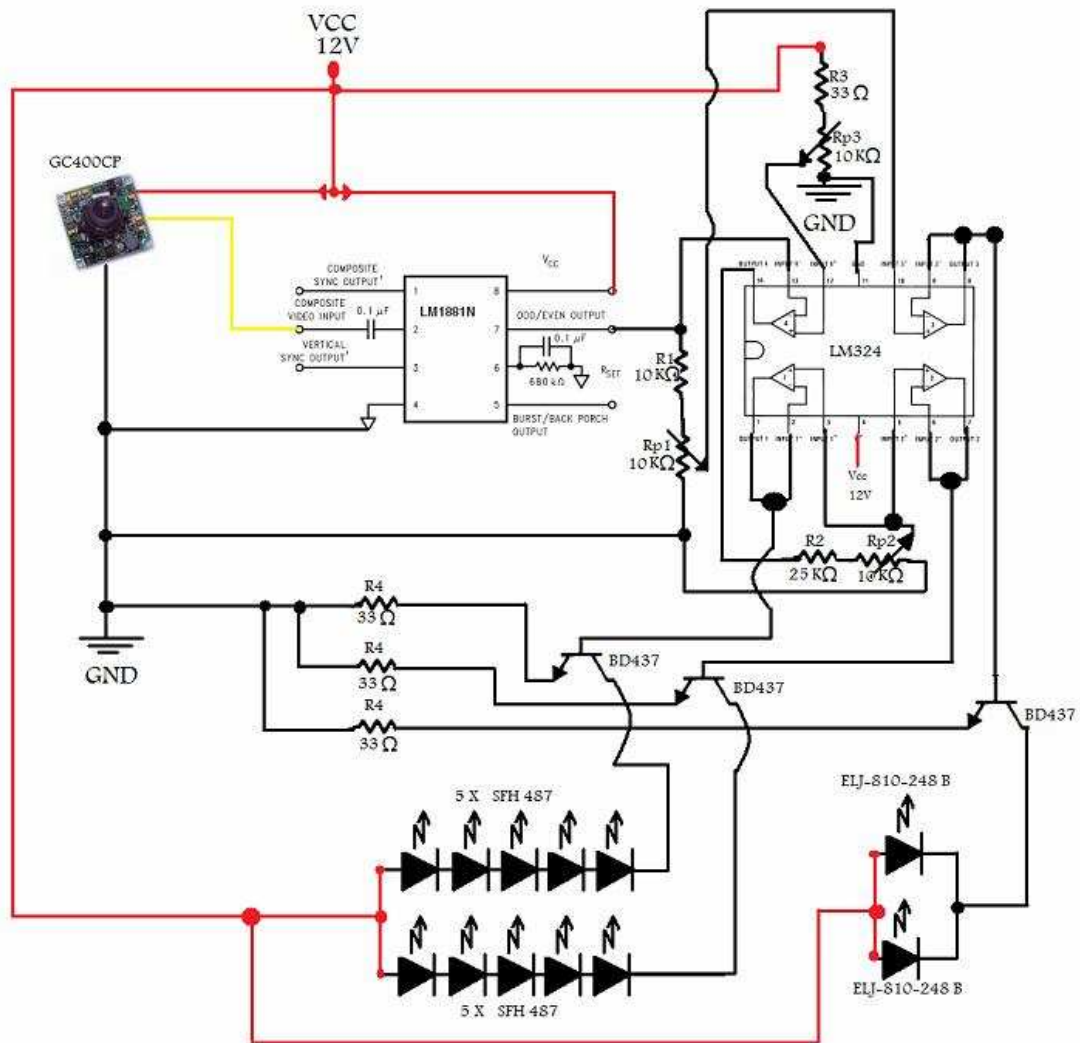
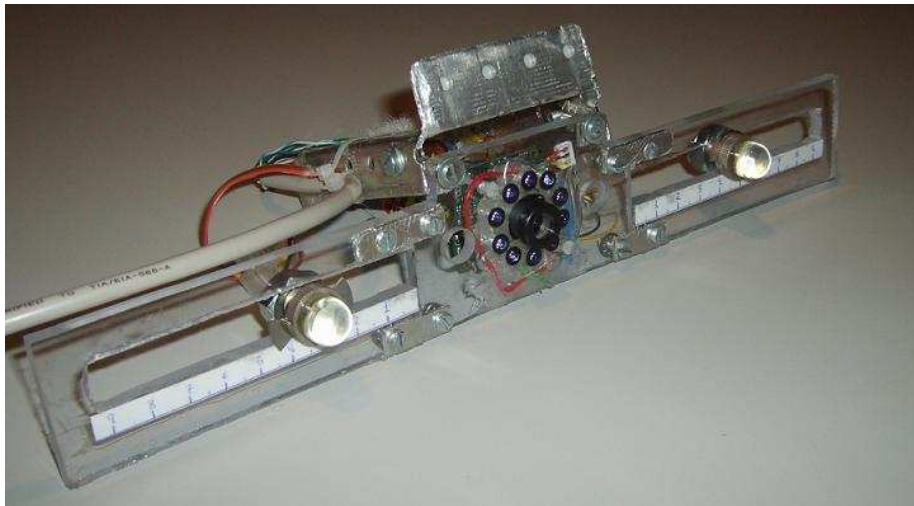
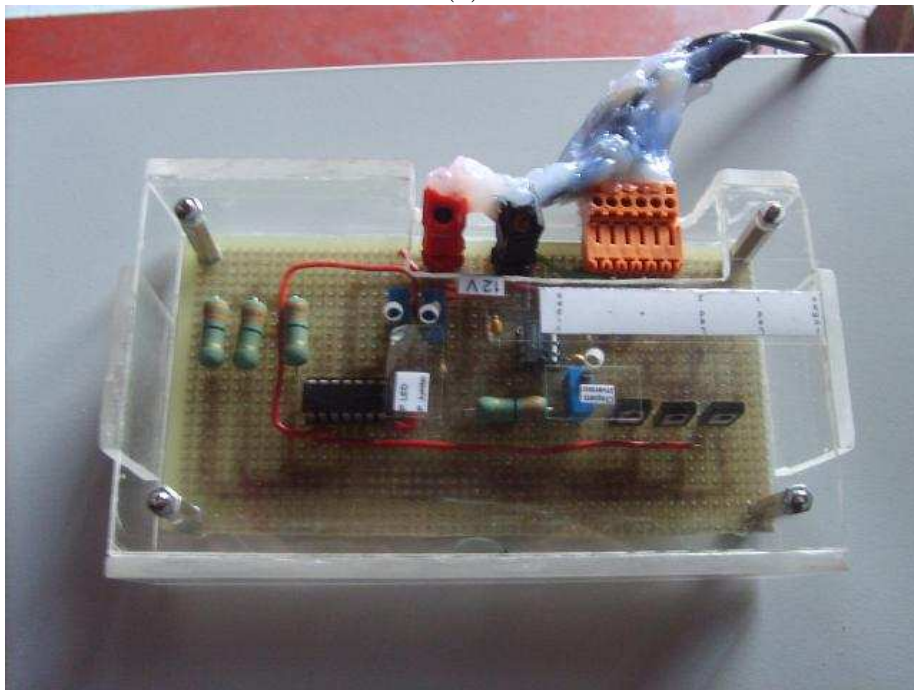


Figura 3.6: Esquema del sistema de visión con iluminación infrarroja.

3.2 Sistema de percepción del vehículo IvviI



(a)



(b)

Figura 3.7: Sistema de visión con iluminación infrarroja y su electrónica asociada, (a) sistema de visión e iluminación, (b) electrónica asociada.



(a)



(b)



(c)

Figura 3.8: Sistema de visión con iluminación infrarroja instalado sobre el vehículo IvvI, (a) sistema encendido, (b), (c) vistas externas.

3.3 Conclusiones



Figura 3.9: Imágenes generadas con el sistema de visión con iluminación infrarroja instalado sobre el vehículo Ivvl.

3.3. Conclusiones

En este capítulo se han presentado tres aspectos importantes para la elaboración de la tesis. El primero corresponde a la identificación de los escenarios de trabajo, que basados en la información de las estadísticas de accidentes de tráfico y los trabajos previos citados en la bibliografía correspondiente, corresponden al día y la noche y que tienen relación directa con la conducción diurna y nocturna. En segundo lugar se ha generado el esquema para la construcción del software en los dos escenarios. Finalmente, el estudio y diseño para la elaboración del hardware ha sido expuesto.

En cuanto al hardware, debido a los escenarios donde se presenta la somnolencia ha existido la necesidad de construir dos sistemas de visión, uno en el espectro visible y otro en el infrarrojo cercano. En el primer caso se utiliza un sistema de visión compuesto principalmente por una cámara, en cambio, en el segundo caso se ha construido un complejo sistema de visión basado en la iluminación infrarroja, específicamente, la que corresponde al infrarrojo cercano. Al final, el primer sistema pasa a formar parte del segundo.

Finalmente, el diseño ha llevado a que se pueda construir un sistema de visión de tamaño reducido, ideal para distintas aplicaciones del tipo análisis del conductor.

Detección de la Somnolencia en el Espectro Visible

Este es uno de los capítulos fundamentales de la investigación, porque contiene en detalle la información necesaria para implementar los algoritmos que componen el sistema de detección de la somnolencia del conductor en el espectro visible, esto quiere decir, que el módulo que se presenta funcionará bajo las condiciones de iluminación natural generadas durante el día, procesando la información proveniente del sistema de visión diseñado para este escenario.

4.1. Descripción del sistema en el espectro visible: diseño y alcance

El sistema propuesto trabaja en varias etapas para garantizar su funcionamiento, generando resultados óptimos frente a las distintas situaciones que se presentan. Algunas de ellas son las que se describen a continuación:

- **Cambios de iluminación:** Los vehículos son utilizados a la interperie en carreteras y autopistas, por lo que se encuentran sometidos a distintos cambios de iluminación, que no es posible controlar.

- **Comportamiento del conductor:** El conductor en general, además de las labores de conducción, realiza otras actividades (manipulación de la radio, atención a las señales de tráfico, entre otras) que interfieren en su desempeño frente al volante.
- **Situaciones imprevistas:** Las imágenes que se pueden obtener del conductor se encuentran corrompidas por varios factores como: gafas de sol, bufandas, gorras, barba, etc. que obstruyen el análisis que se puede realizar sobre el rostro. A esto se le agrega que el vehículo esta en constante movimiento, sometido a vibraciones que generan distorsión en las imágenes.

Todo lo anterior ha sido analizado y tomado en cuenta para el desarrollo e implementación. Por tal motivo, el sistema se encuentra dividido en varias etapas, cada una de ellas garantizan la eficiencia del algoritmo.

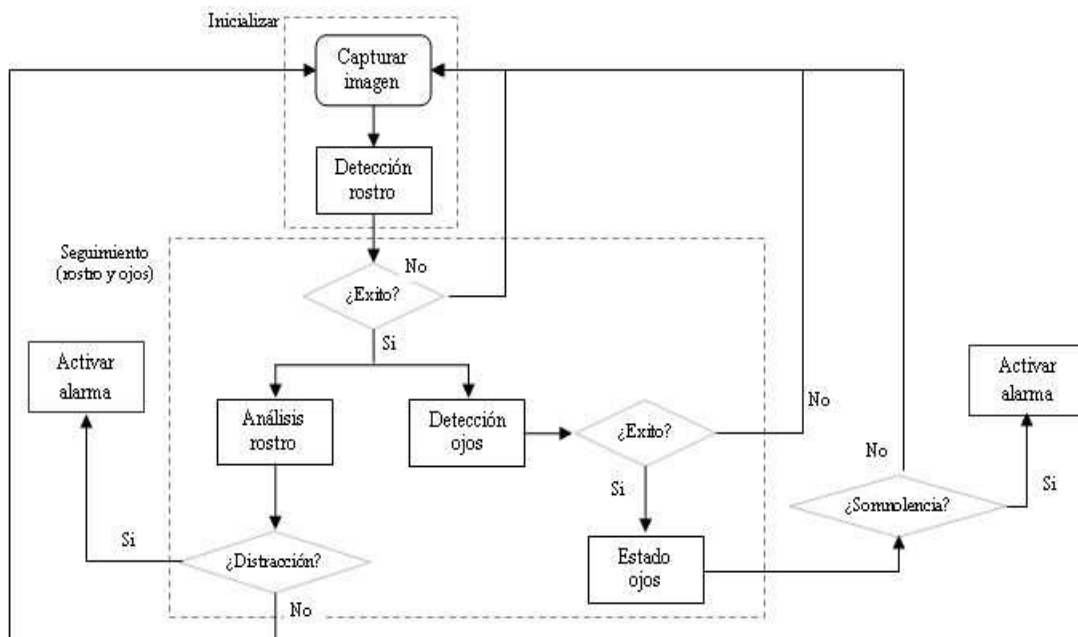


Figura 4.1: Esquema del algoritmo propuesto en el espectro visible.

En la figura 4.1 se presenta el diagrama de trabajo propuesto que se compone de los siguientes módulos:

- Detección del rostro y los ojos

4.2 Detección del rostro del conductor

- Seguimiento del rostro y los ojos
- Cálculo del índice de somnolencia
- Análisis de la distracción

En las próximas secciones se irán describiendo los algoritmos que componen este sistema, para validar las propuestas realizadas se las acompañará de varios ejemplos.

4.2. Detección del rostro del conductor

El rostro y sus características faciales han atraído la atención desde épocas inmemoriales por la cantidad de información visual que pueden proporcionar de una persona (estado de ánimo, género, edad, etc.), en la actualidad, gracias a los avances tecnológicos, su detección y análisis tienen varios campos de aplicación, por ejemplo, se pueden citar los sistemas de identificación de rostros [6], [55], [20], sistemas de vigilancia y seguridad [117], sistemas de tele-conferencias [85], video juegos [53], realidad virtual [65], interacción humano computador [48] [91], biometría [162], etc.

Claramente se concluye que cada una de estas aplicaciones necesitan detectar la posición del rostro, por tal motivo, en el transcurso de los últimos años varios investigadores han planteado distintas aproximaciones para su solución, todo este desarrollo ha venido acompañado del avance en el poder de cálculo de los computadores, ya que sin ellos, el trabajo sería casi imposible.

4.2.1. Clasificación de los métodos para detectar el rostro

En este contexto se han desarrollado varios métodos para la detección del rostro que pueden ser clasificados, según Yang *et al.* [63] y Lee [158], en cuatro grandes grupos:

- **Métodos basados en el conocimiento** (knowledge-based methods): Los métodos basados en el conocimiento utilizan la información previa que se posee sobre los rostros humanos, por ejemplo, medidas antropométricas del rostro, posición del rostro respecto al cuerpo, etc. Utilizan reglas para describir las características de un rostro

y sus relaciones. Las limitaciones que presentan estos métodos tienen que ver con la dificultad de trasladar el conocimiento a reglas bien definidas.

- **Métodos basados en características invariantes** (feature invariant approaches):

Los métodos basados en las características invariantes utilizan, para la detección, ciertos rasgos faciales como ojos, nariz, boca, color y textura; con la información extraída se construyen modelos que describen las relaciones y verifican la existencia de un rostro. Esta clase de algoritmos presentan dificultades cuando las imágenes son corrompidas por la iluminación, ruido y oclusiones.

- **Métodos basados en el emparejamiento de modelos** (template matching methods):

Los métodos basados en el emparejamiento de modelos utilizan modelos predefinidos del rostro y a través del valor de la correlación se establece su existencia. Tienen la desventaja de ser inadecuados cuando existen variaciones de escala, posición y forma.

- **Métodos basados en la forma** (appearance-based methods):

Los métodos basados en la forma trabajan con modelos que se aprenden a partir de ejemplos, como puede ser el caso de imágenes de rostros y no rostros. En general, estos métodos utilizan técnicas estadísticas y de reconocimiento de patrones para encontrar las características relevantes de las imágenes de entrenamiento. Las características aprendidas se presentan en la forma de modelos de probabilidad o en la forma de funciones discriminantes que son debidamente utilizadas para la detección de rostros. Para construir los modelos de probabilidad un vector de características es construido a partir de una imagen y a continuación es tratado como una variable aleatoria. Las funciones discriminantes se definen sobre espacios de baja dimensión utilizando imágenes patrones. Las principales dificultades se encuentran en la dimensión de los vectores de características, que por lo general es alta, esto hace que se originen problemas tanto de cómputo como de interpretación y representación.

4.2 Detección del rostro del conductor

4.2.2. Localización del rostro

En el último caso de la clasificación arriba descrita se encuentra un método que ha generado buenos resultados para detectar el rostro, éste es el algoritmo propuesto por los autores Viola & Jones [120], [118], [119], que toma como base el algoritmo de aprendizaje denominado AdaBoost, presentado por Freund y Schapire en 1.995 [157].

El algoritmo de Viola & Jones genera un rectángulo denominado región de interés (*RI*) que localiza la posición rostro, bajo las siguientes coordenadas:

- Esquina superior izquierda: (x_0, y_0)
- Esquina inferior derecha : (x_1, y_1)

a partir de esta información se obtiene sus dimensiones, ancho y alto, a través de las siguientes ecuaciones:

- Ancho región : $w = x_1 - x_0$
- Alto región : $h = y_1 - y_0$

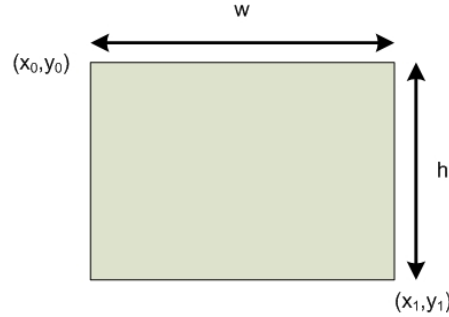


Figura 4.2: Información de la región de interés generada con el algoritmo de Viola & Jones.

De esta manera la *RI* queda completamente definida por las coordenadas de la esquina superior izquierda (x_0, y_0) y las dimensiones del ancho y el alto (w, h) . Esta notación será útil en las siguientes secciones, en especial, para describir el algoritmo de seguimiento.

En la figura 4.2 se observa gráficamente lo descrito en el conjunto de ecuaciones que definen una *RI*. Entretanto en las imágenes (a), (b) y (c) de la figura 4.3 se observa el rectángulo que localiza el rostro de varios conductores sobre el vehículo Ivvi.

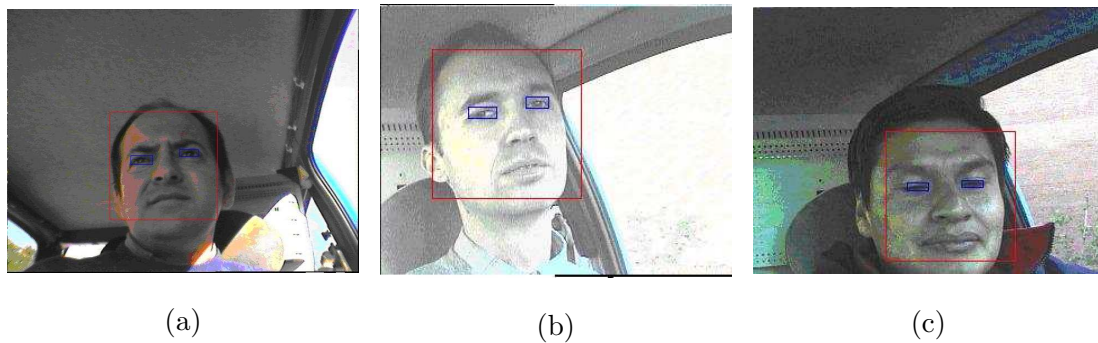


Figura 4.3: Resultados del método propuesto para localizar el rostro y los ojos: (a), (b), (c) localización del rostro con el algoritmo de Viola & Jones.

A pesar de las grandes ventajas que presenta el algoritmo *VJ*, este no es absolutamente fiable para realizar el trabajo de análisis del conductor debido a que no es capaz de localizar el rostro cuando éste no se encuentra dentro de los rangos permitidos en la etapa de entrenamiento. Esto incluye giros laterales, giros hacia arriba o hacia abajo; todos ellos habituales durante la conducción y que se producen en situaciones como el mirar el espejo retrovisor y/o espejos laterales, etc.

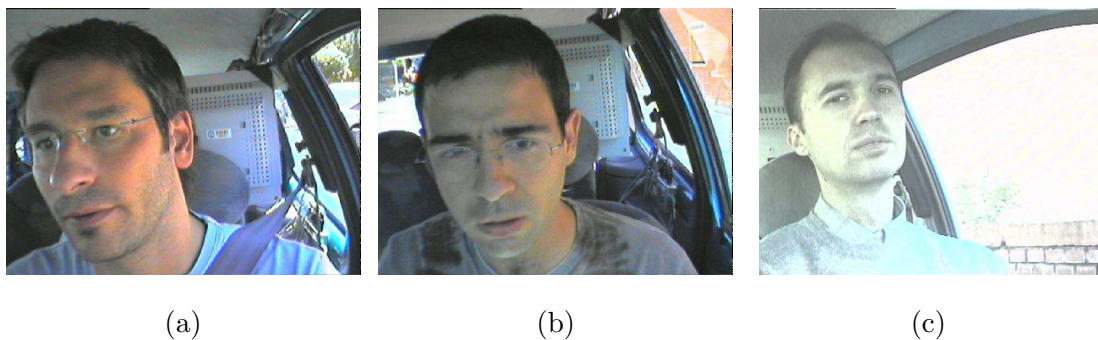


Figura 4.4: Varios ejemplos donde el algoritmo de Viola & Jones no detecta la presencia del rostro.

Esta dificultad será analizada y corregida más adelante, hasta ello, en la figura 4.4 se exhiben varias imágenes que contienen un rostro que no puede ser detectado por el algoritmo de VJ.

4.3. Detección de los ojos

Los ojos constituyen la parte fundamental del sistema de detección de la somnolencia, motivo por el cual, se hace hincapié en la revisión de los trabajos previos, antes de presentar el método propuesto en la presente tesis.

4.3.1. Trabajos previos para detectar los ojos

La detección y análisis de los ojos es vital para diseñar diferentes aplicaciones humano-computador; en este contexto, el campo de aplicaciones es extenso y va desde la medicina hasta la publicidad [121], por tal motivo, para detectar los ojos se han propuesto varios trabajos utilizando diferentes técnicas, por ejemplo, Wang y Yang [126] presentan un método que detecta los ojos sobre una imagen en escala de grises utilizando umbralización simple bajo la hipótesis de que los ojos corresponden a los píxeles de menor valor en la escala de grises (aproximadamente corresponderían a los píxeles que forman el color negro) para construir una imagen binaria donde se localizan varios cuerpos, entre ellos los ojos, luego, se discrimina entre ojos y no-ojos utilizando el método de clasificación SVM. Liu *et al.* [62] a su vez propone detectar los ojos a través de la función normal bivalente que modela el color de la piel sobre el espacio normalizado RGB; bajo esta técnica, la zona que corresponde a los ojos será la más oscura dentro de la zona del rostro. Una vez localizados los ojos se verifica su estado, para ello se utiliza la información del iris y la esclera en el espacio de color HSI, para verificar si está abierto o cerrado. Siguiendo este método se encuentra el trabajo de Kashima *et al.* [61] sobre el espacio de color CIE-LUV.

4.3.2. Inconvenientes para detectar los ojos

Sin embargo, a pesar de todos los esfuerzos descritos, los métodos propuestos presentan limitaciones ya que trabajan bajo condiciones constantes de iluminación como son las condiciones de laboratorio; a esto hay que añadir otras limitaciones que presentan la mayoría de sistemas de seguimiento y localización de los ojos, que es, que ellos trabajan sobre los ojos abiertos y presentan muchos falsos positivos cuando los ojos se encuentran cerrados [160].

Localizar la posición exacta de los ojos es un trabajo altamente complicado porque involucra muchos parámetros que no permiten construir un método lo suficientemente robusto. Entre los principales inconvenientes que se pueden citar se encuentran:

- las diferentes formas que adopta el ojo, debido principalmente a dos factores, la necesidad biológica de parpadear y el movimiento del rostro,
- los cambios de iluminación, en especial en ambientes al aire libre, como es el caso de la conducción, y
- una característica propia de los ojos, el color del iris (incluye a la pupila), que dependiendo de la raza, toma varios colores como el marrón, negro, azul, etc.

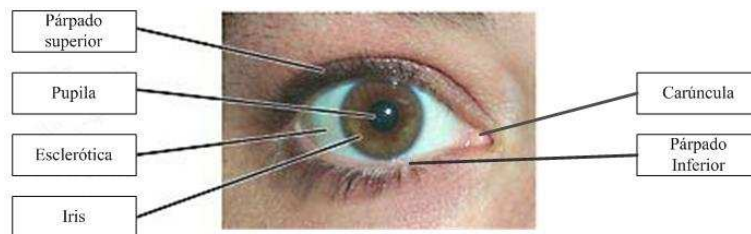


Figura 4.5: Anatomía del ojo.

Todo esto ha llevado a que en esta tesis se busque un método alternativo y altamente fiable para localizarlos, el mismo que se presenta en la sección que viene a continuación.

4.3.3. Método propuesto para detectar los ojos basado en el algoritmo EM

Tomando en cuenta los factores mencionados se procede a describir el método propuesto en este trabajo. La idea principal está inspirada en la teoría de Probabilidades y toma en consideración la geometría del ojo.

Recordar que bajo condiciones de iluminación a la interperie, el valor de los píxeles que forman parte del ojo (en general de una imagen cualquiera) toman valores aleatorios en el intervalo $[0, 255]$. Por lo tanto, con el objetivo de extraer la mayor cantidad de información, proveniente de la geometría del ojo bajo iluminación cambiante, se plantea el algoritmo

4.3 Detección de los ojos

para su detección. En este contexto, a priori se conoce que la forma del ojo es elíptica, más específicamente, el ojo puede ser visto como una realización de una función de probabilidad bidimensional; por ejemplo, la distribución normal bivariada o una distribución elíptica [56], [127], [92]. Con esta información a priori se desarrolla el algoritmo que se propone, el mismo que se basa en una serie de imágenes que resaltan las características geométricas del ojo, visto éste, como una elipse.

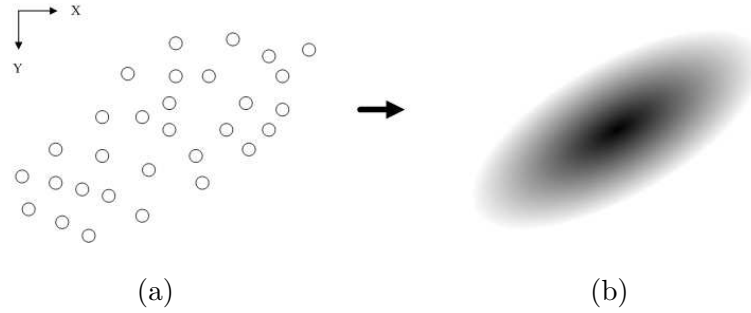


Figura 4.6: Muestra aleatoria de la forma del ojo: (a) realización de una función de probabilidad, (b) modelo ajustado a partir de la muestra.

Para aclarar la idea, en la figura 4.6 (a) se puede ver una serie de puntos agrupados que forman una elipse, esto es lo que se denominará una muestra. En la misma figura, en la parte (b), se observa una elipse que ha sido ajustada usando la muestra de la parte (a), este modelo representa y reproduce con bastante precisión el fenómeno que genera la muestra. Este breve ejemplo describe la idea del método propuesto para detectar los ojos; por otra parte, para ajustar dicho modelo, más adelante se plantea una solución basada en el algoritmo EM [2], [84].

Tomando como punto de partida la idea anterior, para detectar los ojos es necesario extraer una muestra representativa que genere la suficiente información de lo que es un ojo. Para ello, usando una serie de técnicas, que se describirán a continuación, se logra extraer ésta muestra aleatoria.

Las coordenadas iniciales de los ojos vienen dadas por las coordenadas de los rectángulos RI_L para el ojo izquierdo y RI_R para el ojo derecho. Sus coordenadas se calculan a

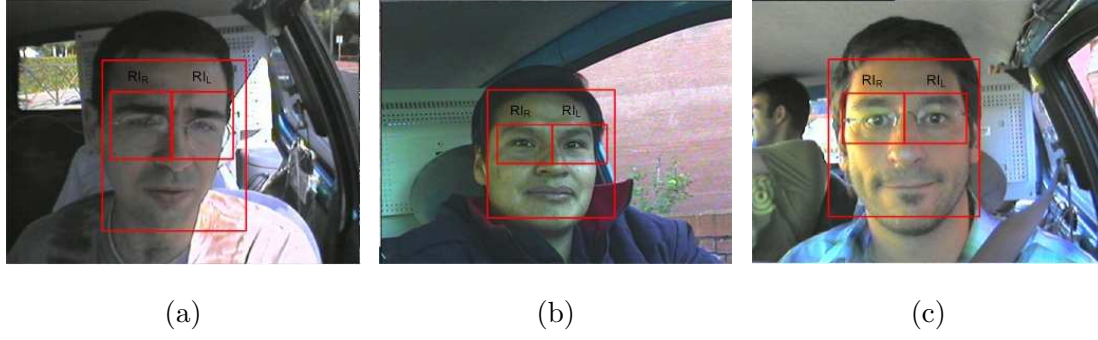


Figura 4.7: Rectángulos preliminares para detectar los ojos.

través de las ecuaciones (4.1), (4.2), (4.3) y (4.4), respectivamente. En la figura 4.7 se aprecian ejemplos de dichos rectángulos sobre varios conductores.

$$(u_{0L}, v_{0L}) = (x_0 + \frac{w}{6}, y_0 + \frac{h}{4}) \quad (4.1)$$

$$(u_{1L}, v_{1L}) = (x_0 + \frac{w}{2}, y_0 + \frac{h}{2}) \quad (4.2)$$

$$(u_{0R}, v_{0R}) = (x_0 + \frac{w}{2}, y_0 + \frac{h}{4}) \quad (4.3)$$

$$(u_{1R}, v_{1R}) = (x_1 - \frac{w}{6}, y_1 - \frac{h}{2}) \quad (4.4)$$

Para extraer la muestra aleatoria de píxeles que forman parte del ojo se considera el siguiente algoritmo, donde se resaltan los píxeles que forman el ojo y que le dan la característica intrínseca respecto a los otros elementos del rostro. La muestra se formará con la información proveniente de tres imágenes que se denominarán B, G y L.

Sea I la imagen de entrada e $I(x, y)$ el valor del pixel en la escala de grises, es decir, $I(x, y) \in [0, 255]$ en el punto de coordenadas (x, y) , entonces a partir de aquí se tiene:

1. Generar la imagen J a través de la ecuación:

$$J(x, y) = \frac{I(x, y) - m}{\sigma} \quad (4.5)$$

4.3 Detección de los ojos

donde m es la media y σ es la desviación estandar, estos valores son calculados sobre los rectángulos descritos anteriormente.

2. Generar la imagen H usando la ecuación (4.6):

$$H(x, y) = \begin{cases} J(x, y) - 256\delta_1 & \text{si } J(x, y) \geq 0 \\ 256\delta_2 + J(x, y) & \text{si } J(x, y) < 0 \end{cases} \quad (4.6)$$

donde $\delta_1 = \max(0, \text{ceil}(\frac{J(x,y)}{256}) - 1)$, $\delta_2 = \max(1, \text{ceil}(\frac{|J(x,y)|}{256}))$ y $\text{ceil}(t)$ es la función que retorna el menor entero mayor que t . La validez de las variables δ_1 y δ_2 queda justificada porque los valores de la imagen (matriz) $J(x, y)$ son números reales. El intervalo para δ_1 es $\{0, 1, \dots\}$ mientras que para δ_2 es $\{1, 2, \dots\}$. Varios ejemplos se presentan en la figura 4.8.

3. Generar la imagen binarizada B , a partir de la imagen H , a través de la siguiente ecuación:

$$B(x, y) = \begin{cases} 255 & \text{si } H(x, y) \geq \kappa \\ 0 & \text{caso contrario} \end{cases} \quad (4.7)$$

donde el umbral κ se calcula automáticamente con el método de Otsu [107], en la figura 4.9 (b) se muestra el resultado.

4. Calcular la imagen gradiente, G , utilizando el operador de Sobel, seguido de un realce de bordes, figura 4.9 (c).

$$G = \sqrt{\|I * S_x\|^2 + \|I * S_y\|^2} \quad (4.8)$$

donde $\|\cdot\|$ es la norma euclídea y las máscaras son:

$$S_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}, \quad S_y = -S_x^T$$

donde T representa la transpuesta de una matriz.

5. Calcular la imagen logaritmo, L , con el objetivo de realzar los píxeles que corresponden al iris y que constituyen la parte central del ojo [161], figura 4.9 (d).

$$L(x, y) = \log(1 + I(x, y)) \quad (4.9)$$

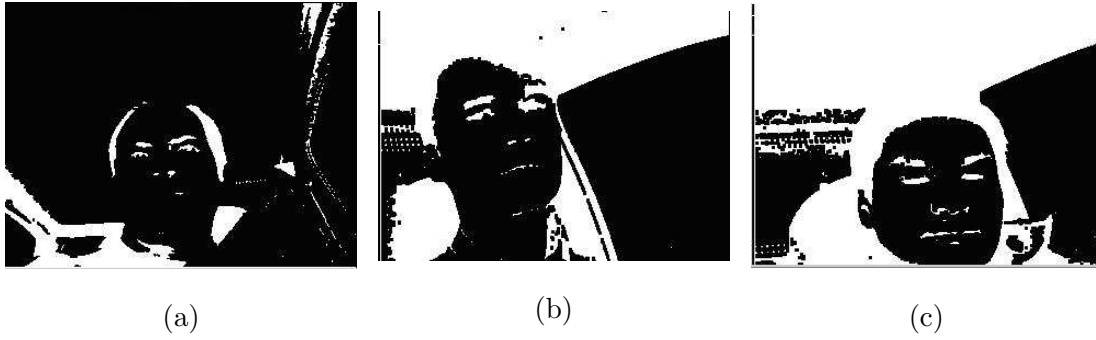


Figura 4.8: Resultados de la segmentación de los ojos, imagen H: (a), (b) y (c).

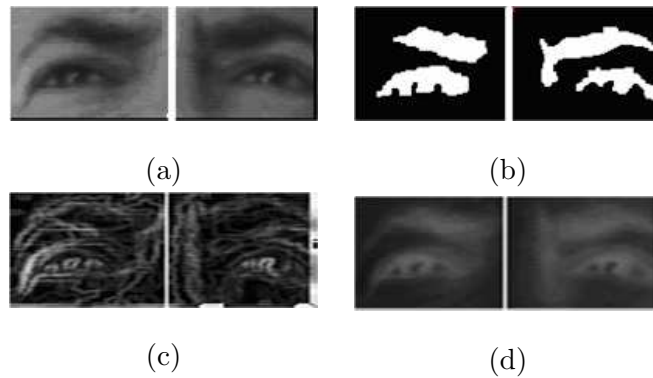


Figura 4.9: Imágenes utilizadas para localizar la posición de los ojos: (a) imagen original, (b) imagen B , (c) imagen G y (d) imagen L .

Con la muestra aleatoria obtenida y con la información a priori de la forma del ojo es posible emplear técnicas paramétricas para estimar los parámetros que generan su forma. Esto quiere decir, que en este caso para localizar la posición del ojo se puede acudir a algún método estadístico que permita estimar los parámetros del modelo elíptico, mencionado anteriormente.

En este contexto, el algoritmo EM^1 (Expectation Maximization) permite calcular los parámetros del modelo [84]. Para ejecutar el algoritmo EM se precisa de un punto inicial (parámetros de inicialización) para empezar el proceso de iteración, este punto inicial se calcula con el conocido método de mínimos cuadrados. Para cumplir las restricciones de tiempo real se permite un número acotado de iteraciones, en este caso particular se

¹Ver el apéndice para conocer los detalles de este algoritmo

4.3 Detección de los ojos

consigue con 10 iteraciones sobre una muestra en dos dimensiones y una gaussiana.

Al trabajar con la muestra aleatoria proveniente de la zona del ojo por medio de tres fuentes distintas se puede tener la certeza, que al menos una de las imágenes (B , G o L) generarán la información necesaria para ejecutar el algoritmo EM .

Para finalizar esta sección se presentarán varios ejemplos que validan la propuesta, en la figura 4.10 (c) se muestran los ejes mayor y menor de la eplise del ojo y el ángulo que forman respecto al eje de coordenadas, representado en la figura 4.10 (a). Mientras que en la figura 4.11 se observan más ejemplos desarrollados sobre varios conductores.

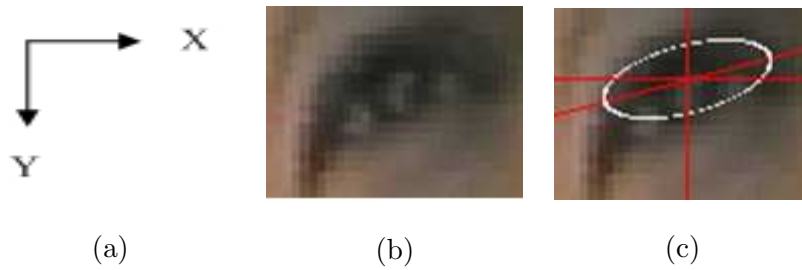


Figura 4.10: Resultado del algoritmo EM al modelar la forma del ojo: (a) sistema de coordenadas, (b) imagen original, (c) elipse ajustada sobre el ojo.

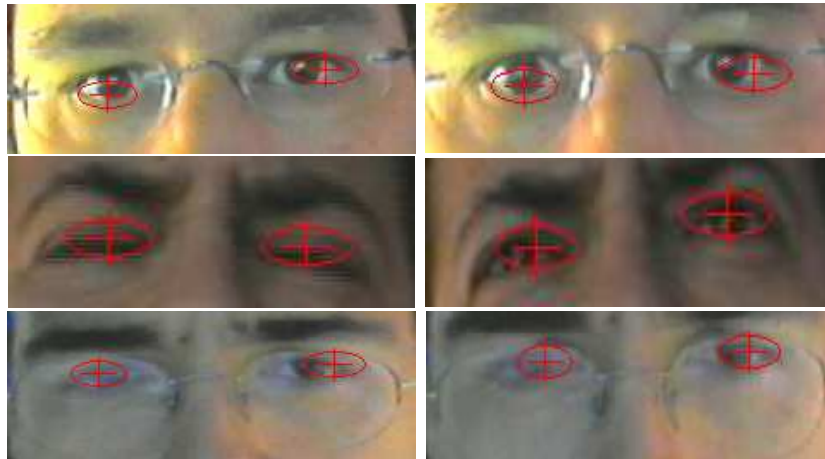


Figura 4.11: Varios ejemplos donde se localiza la posición de los ojos con el método propuesto.

4.4. Enfoque bayesiano para el seguimiento del rostro y los ojos

4.4.1. Motivos para el seguimiento

Existen varias razones para desarrollar el seguimiento, entre ellas se tienen las siguientes:

- El método de VJ a pesar su gran adaptación de trabajo en ambientes con iluminación variable, presenta ciertos inconvenientes que generan inestabilidad en el sistema y lo inducen a cometer errores. En la figura 4.12 se presenta un ejemplo extendido donde el algoritmo VJ no detecta el rostro sobre una secuencia de imágenes donde claramente se encuentra uno.



Figura 4.12: El rostro del conductor no es detectado por el método de Viola & Jones en varios instantes de tiempo.

- Otra es la necesidad de conocer la posición del rostro y los ojos continuamente en cada instante, es decir, en cada fotograma.
- La tercera razón es satisfacer las condiciones de tiempo real, reduciendo el espacio de búsqueda y en lo posible disminuyendo el tiempo de cómputo.

Para desarrollar algoritmos de seguimiento se han propuesto varias estrategias, entre ellas utilizar el filtro de Kalman [133], [156], y recientemente en el campo de la visión por computador el algoritmo de Condensation [93], [94], que utiliza el filtro de partículas [3],

[10] para evaluar las complejas funciones de densidad de probabilidad que se generan en el modelo probabilístico que evoluciona a través del tiempo.

4.4.2. Descripción del algoritmo Condensation

Para el seguimiento del rostro y los ojos se utiliza el algoritmo de *Propagación de la Densidad Condicional*, Condensation, desarrollado por Michael Isard y Andrew Blake en 1.998 para el seguimiento de contornos activos sobre secuencias de imágenes [93], [94], [5].

El algoritmo Condensation combina un método de muestreo denominado factored sampling (un método Monte-Carlo [57]) con un modelo dinámico que es gobernado por la ecuación estocástica (4.10), que describe la evolución del sistema.

$$x_t = f(x_{t-1}, \xi_t) \quad (4.10)$$

donde x_t es el estado en el instante t , $f(\cdot)$ es una función no lineal que depende del estado previo más un ruido blanco ξ_t .

En los problemas de seguimiento, el objetivo es estimar el estado futuro x_t utilizando la información de las observaciones realizadas sobre el sistema, estas observaciones se expresan a través de la ecuación (4.11).

$$z_t = h(x_t, \eta_t) \quad (4.11)$$

donde z_t es la medida realizada sobre el sistema al instante t , $h(\cdot)$ es otra función no lineal que vincula la medida con el estado presente más el ruido blanco η_t . Los procesos ξ_t y η_t son procesos independientes entre sí. En general, estos procesos son no gaussianos y multi modales.

Bajo los supuestos de proceso de Markov para el estado se tiene la función de probabilidad de la dinámica del sistema.

$$p(x_t) = \int p(x_t|x_{t-1})p(x_{t-1})dx_{t-1} \quad (4.12)$$

donde $p(x_t|x_{t-1})$ es la densidad del proceso. Esta ecuación es utilizada para estimar la distribución de probabilidad en el siguiente instante de tiempo.

En el problema de seguimiento se desea estimar el estado en cada instante incorporando la información de la medida z_t . Se incorpora la función de densidad a priori $p(x_t|z_{1:t-1})$ y la función de densidad a posteriori $p(x_t|z_{1:t})$, donde $z_{1:t} = \{z_1, \dots, z_t\}$ representa la historia de las observaciones hasta el instante t .

Usando el teorema de Bayes se determina la densidad a posteriori $p(x_t|z_{1:t}) = p(x_t|z_t, z_{1:t-1})$ a partir de la densidad a priori.

$$\begin{aligned}
 p(x_t|z_{1:t}) &= \frac{p(z_t|x_t, z_{1:t-1})p(x_t|z_{1:t-1})}{p(z_t|z_{1:t-1})} \\
 &= kp(z_t|x_t, z_{1:t-1})p(x_t|z_{1:t-1}) \\
 &= kp(z_t|x_t)p(x_t|z_{1:t-1})
 \end{aligned} \tag{4.13}$$

donde k es una constante de normalización. Las simplificaciones se justifican bajo el supuesto de independencia de las medidas. La función de densidad de las observaciones $p(z_t|x_t)$ define la probabilidad de que el estado x_t cause la medida z_t .

La parte más complicada de evaluar en el algoritmo Condensation es estimar la probabilidad condicional a posteriori [138], [42], lo que conlleva a evaluar la función de densidad de las observaciones $p(z_t|x_t)$, que cuantifica la relación entre el estado y las medidas realizadas sobre el sistema.

Para ello, en el campo de la visión por computador se han generado varias estrategias, entre ellas se encuentran, por ejemplo, la correlación [75], [21], los modelos de apariencia activa (Active Appearance Models) [134] y los modelos de contorno formables [5], [93], [94], los snakes [22], la probabilidad asociada a los modelos probabilísticos del color [138], [27], etc.

En la estructura del algoritmo Condensation, la densidad a posteriori $p(x_t|z_{1:t})$ es representada por una muestra (partículas) $\{s_t^{(n)}, n = 1, \dots, N\}$ y la densidad de las observaciones dado el estado es aproximada por los pesos $\{\pi_t^{(n)}, n = 1, \dots, N\}$. Con esto se tiene la aproximación:

4.5 Seguimiento del rostro

$$\pi_t^{(n)} \approx p(z_t | x_t = s_t^{(n)}) \quad \text{para } n = 1, \dots, N \quad (4.14)$$

que será utilizada en las siguientes secciones.

4.5. Seguimiento del rostro

En esta sección se describe en detalle el algoritmo de seguimiento del rostro; el algoritmo propuesto incluye la información generada por dos fuentes, el algoritmo Condensation y las redes neuronales artificiales.

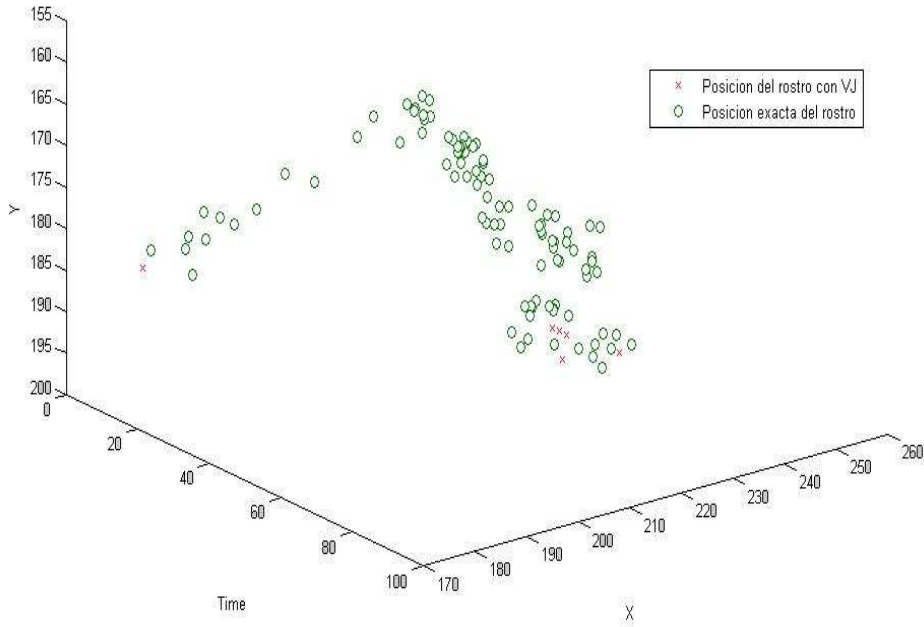


Figura 4.13: Comparación entre la posición exacta y la posición generada por VJ sobre una secuencia de 100 imágenes.

En complemento a la figura 4.12, se presenta la figura 4.13 que indica la trayectoria generada por el algoritmo de VJ sobre una secuencia compuesta por 100 imágenes de un conductor sobre el vehículo IvvI. En esta secuencia la tasa de detección se encuentra en el orden del 13%. Esta baja tasa de detección constituye un serio problema en este tipo de sistemas, donde es imprescindible mantener localizado el rostro en todo instante. Si el

vehículo circula a 80 km/h, esto significa que ha recorrido aproximadamente 88,8 metros sin que el sistema sea capaz de detectar el rostro en condiciones de tiempo real, es decir, a la tasa de 25 imágenes por segundo.

4.5.1. Especificación del modelo de seguimiento del rostro

El modelo dinámico del sistema para el seguimiento del rostro es un modelo autoregresivo de segundo orden² $AR(2)$, que se representa de acuerdo a la ecuación (4.15).

$$x_t = A_2 x_{t-2} + A_1 x_{t-1} + \xi_t \quad (4.15)$$

donde A_2 y A_1 relacionan el estado presente con los estados anteriores y son similares a las descritas por Blake [5] y Doucet [3], ξ_t es un ruido blanco .

El modelo de observación viene dado por la ecuación (4.16).

$$z_t = H x_t + \eta_t \quad (4.16)$$

donde H relaciona el estado con las observaciones actuales, su estructura de detalle viene en Doucet [3], y η_t es un ruido blanco.

A continuación se definen los parámetros del vector de estados que gobierna al sistema dinámico, en este caso particular, se desea conocer la posición del rostro dentro de la imagen, su velocidad de desplazamiento; y por supuesto, las dimensiones que ocupa el rectángulo que lo contiene. Por otra parte está el vector de medidas, que se compone de las medidas que se pueden extraer directamente de la imagen, en este caso, se compone de la posición y las dimensiones del ancho y alto del rectángulo que contiene al rostro.

Una vez que se han descrito las variables que componen el sistema dinámico (vector de estado y vector de medidas) se procede a escribirlas en lenguaje matemático. El estado del rostro, en el estado t , viene dado por la ecuación (4.17).

$$x_t = (x, y, u, v, w, h) \quad (4.17)$$

²Second Order Autoregressive Process (ARP) en inglés

4.5 Seguimiento del rostro

donde (x, y) es la posición del centro del rostro dentro de la imagen, (u, v) es su velocidad y w y h constituyen el ancho y el alto del rostro. La medida, en el instante t , viene dada por la ecuación (4.18):

$$z_t = (x, y, w, h) \quad (4.18)$$

Esto implica que $x_t \in \mathbb{R}^6$ y $z_t \in \mathbb{R}^4$ durante todo el proceso de seguimiento, una vez definidos el vector de estado, el vector de medida y la ecuación del modelo dinámico y el modelo de observación, solo resta evaluar la densidad de las observaciones dado un estado (partícula).

Una vez planteado el modelo de seguimiento del rostro bajo el esquema del algoritmo Condensation, resta introducir un método para evaluar la función de densidad de las observaciones. En este caso concreto su estimación dependerá del valor generado por una red neuronal artificial (RNA) al ser evaluada sobre cada una de las partículas.

4.5.2. Descripción de las redes neuronales artificiales

Las redes neuronales artificiales consisten en técnicas que tratan de imitar el funcionamiento del cerebro humano y a través de ello “aprenden”, teniendo la capacidad de extrapolar y generalizar, motivo por el cual, obtienen excelentes resultados al ser utilizadas en problemas complejos.

Existen varias razones para emplear una RNA, pero se mencionarán las más importantes ya que provienen de las múltiples dificultades encontradas durante el desarrollo de la investigación:

- Cabe recordar que en la mayoría de casos, el rostro no se encuentra plenamente en posición frontal debido a las múltiples actividades desarrolladas para conducir, esto genera que el método de VJ no sea capaz de detectarlo adecuadamente, lo que introduce una importante fuente de errores, lo que impide el normal funcionamiento del sistema.
- Verificar que se trata de un rostro, realizando un barrido completo sobre una vecindad a la zona predicha por el algoritmo Condensation. El rostro en general se encuentra

en posición vertical, pero en muchos casos éste aparece girado, lo que dificulta su detección.

- Mantener al sistema inmune a los cambios de iluminación tomando en cuenta las virtudes de las RNA que constituyen una base de conocimiento y tienen la capacidad de extrapolación y generalización.
- Generar el índice de distracción, que será analizado en la última sección de este capítulo.

4.5.2.1. Definición: Espacio imagen

Para las siguientes secciones será útil una transformación de la imagen, sea I una imagen que tiene representación $2D$ (figura 4.14 (a)) se la transforma en un vector (figura 4.14 (b)), esto es lo que Vicente [7] define como transformación al espacio imagen.

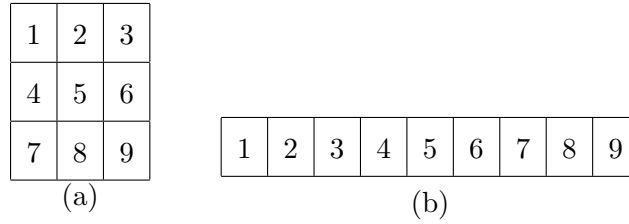


Figura 4.14: Transformación al espacio imagen, (a) representación $2D$, (b) representación $1D$.

4.5.3. Entrenamiento de la red neuronal artificial

El rostro presenta cierta información intrínseca que lo caracteriza y lo diferencia de otras cosas u objetos. Con el ánimo de resaltarlas se incluye la fase de preprocesamiento, antes de entrar de lleno en la fase de entrenamiento.

4.5.3.1. Preprocesamiento y extracción de características

Antes de extraer las características que describen al rostro, que serán las entradas de la RNA, es preciso realizar un paso previo que consiste en adecuar la imagen mediante

4.5 Seguimiento del rostro

diferentes técnicas del procesamiento de imágenes. En este caso se realizan los siguientes pasos:

1. Corregir la iluminación de la imagen mediante la modificación del contraste [28], utilizando en este caso la corrección gamma [131], el método consiste en tomar un pixel de entrada $I(x, y)$ y elevarlo a la potencia γ para obtener un nuevo pixel $J(x, y)$:

$$J(x, y) = I(x, y)^\gamma \quad (4.19)$$

donde γ toma el valor de 0,8 en este caso en particular y se lo ha elegido experimentalmente.

2. Ecualizar el histograma de la imagen obtenida en el paso anterior con la función uniforme [0, 255].
3. Retirar los píxeles de las esquinas mediante la operación AND con la máscara de la figura 4.16 (b).

En la figura 4.15 se presenta un ejemplo de la etapa de preproceso desarrollada sobre un rostro, en fila superior se muestran los rostros y en la fila inferior sus respectivos histogramas, donde se observa la variación que han sufrido bajo el proceso del paso uno y dos.

Una vez obtenida la imagen del rostro se extrae el vector de características que será la entrada de la RNA. En este caso concreto, este vector, corresponde al nivel de grises tomados directamente de la imagen y constituye una representación tan buena como otras [141].

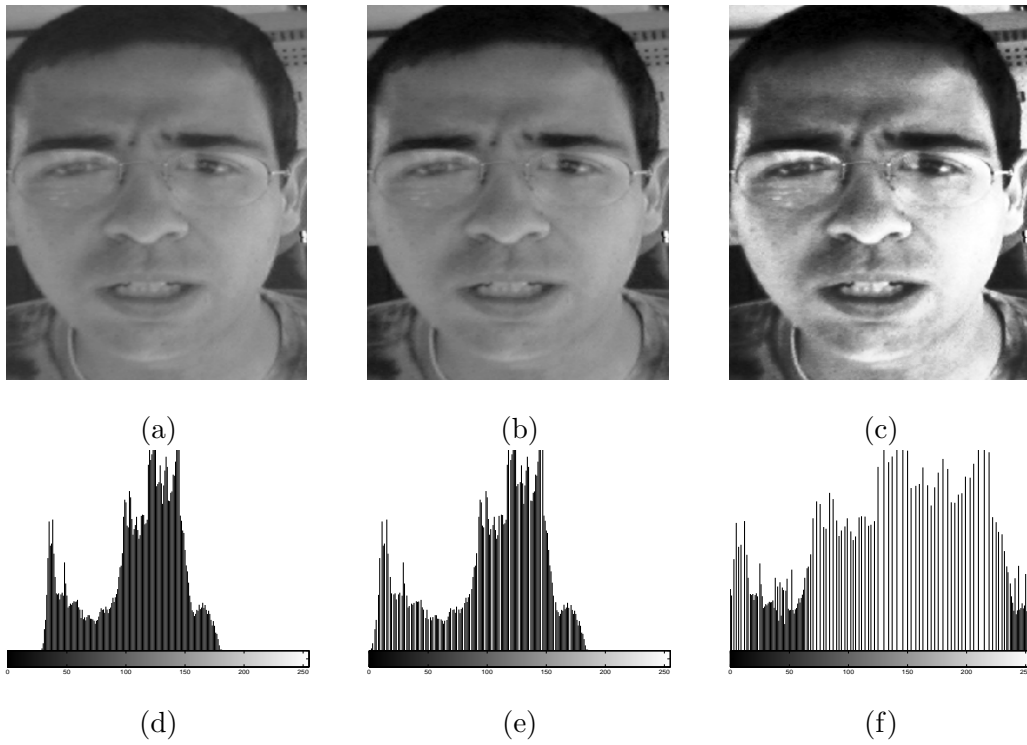


Figura 4.15: Corrección de la iluminación con el factor gamma y ecualización del histograma, (a) imagen original, (b) imagen corregida, (c) imagen ecualizada y (d), (e), (f) sus respectivos histogramas.

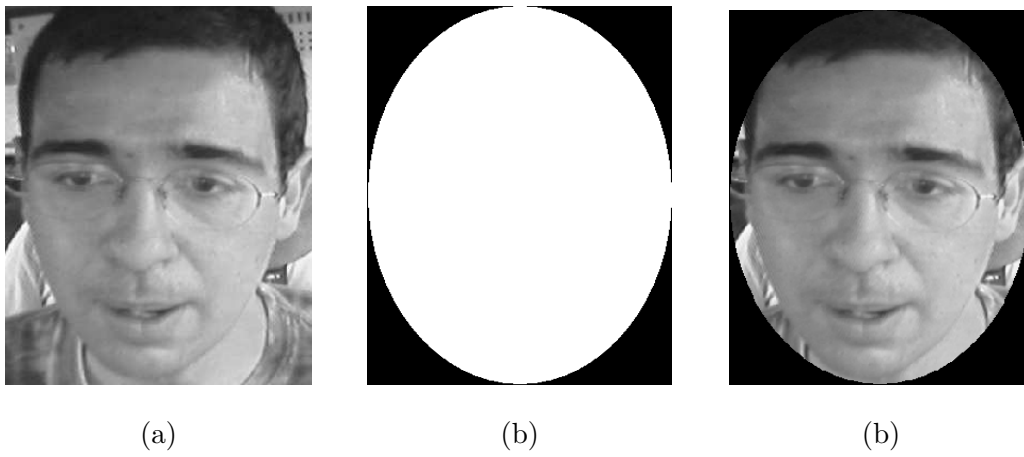


Figura 4.16: Eliminación de las esquinas y operación AND, (a) Imagen de entrada, (b) máscara para remover las esquinas, (b) resultado al remover las esquinas con la operación AND.

4.5 Seguimiento del rostro

4.5.3.2. Entrenamiento

Para entrenar la red neuronal se ha construido una base de datos con rostros y no rostros. Varios de los rostros provienen de algunos conductores del vehículo Ivvi y otros proceden de fotos, internet, etc. Cada elemento de la base pasa por el preprocesamiento para posteriormente convertirlo en un vector. Las dimensiones de cada imagen son de 18 píxeles de ancho por 27 píxeles de alto, a partir de la cual se han generado otras tres, mediante rotaciones de 90° , 180° y 270° , respectivamente.

En la figura 4.17 se presenta un ejemplo de las muestras recolectadas para la tesis, que en total suman 646 rostros y 600 no-rostros. Como se puede ver, los rostros se encuentran sometidos a distintos tipos de iluminación y varias orientaciones. En este caso se ha utilizado 1 para representar un rostro y -1 para representar un no-rostro.

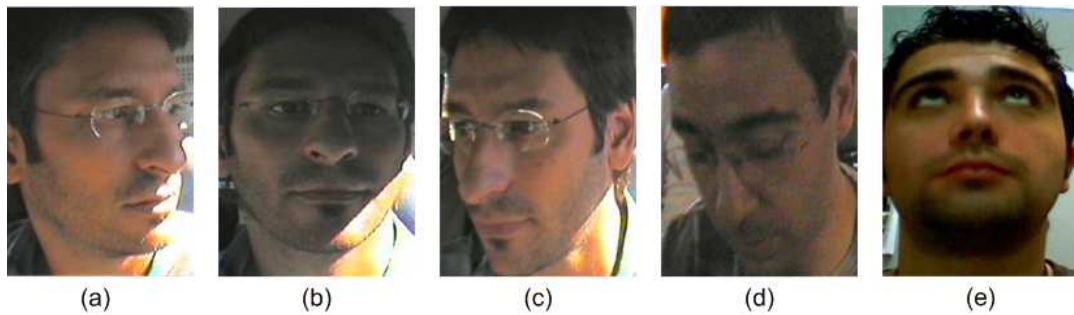


Figura 4.17: Ejemplos de la base de datos de rostros.

El modelo de RNA utilizado es *Backpropagation*, que es una de las más conocidas y utilizadas en problemas de clasificación, sus parámetros de entrenamiento después de varios experimentos son los siguientes:

- Número de neuronas de entrada : 1.944
- Número de neuronas de salida: 1
- Número de capas: 3
- Número de neuronas ocultas: 96
- Error deseado: 1×10^{-4}

- Función de activación: sigmoid

$$f(t) = \frac{1}{1 + e^{-t}} \quad -\infty < t < \infty \quad (4.20)$$

En la fase de prueba para comprobar la eficacia del entrenamiento se ha utilizado una base de datos con rostros y no-rostros que no se encontraban en la base de datos de entrenamiento. Con estos parámetros se ha logrado una tasa de clasificación de alrededor del 93 % sobre 2.000 imágenes de rostros y 2.500 imágenes de no-rostros. La curva ROC (Receiver Operator Curve) del conjunto de prueba se observa en la figura 4.18, que demuestra que existe un número aceptable de falsas detecciones.

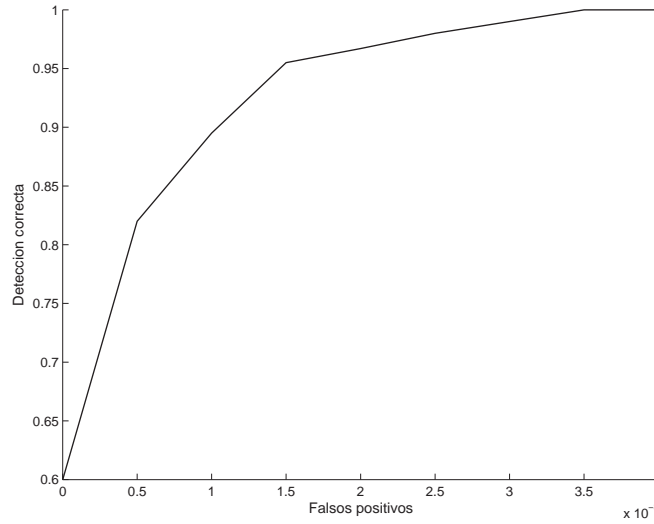


Figura 4.18: Curva ROC generada sobre la base de datos de prueba de rostros.

4.5.4. Evaluación de la función de densidad

En este caso, las partículas se representan como las denominadas regiones de interés (RI) sobre la imagen, un ejemplo se observa en la figura 4.19 (b), las mismas que generan la región predicha de la figura 4.19 (a) que enmarca al rostro del conductor.

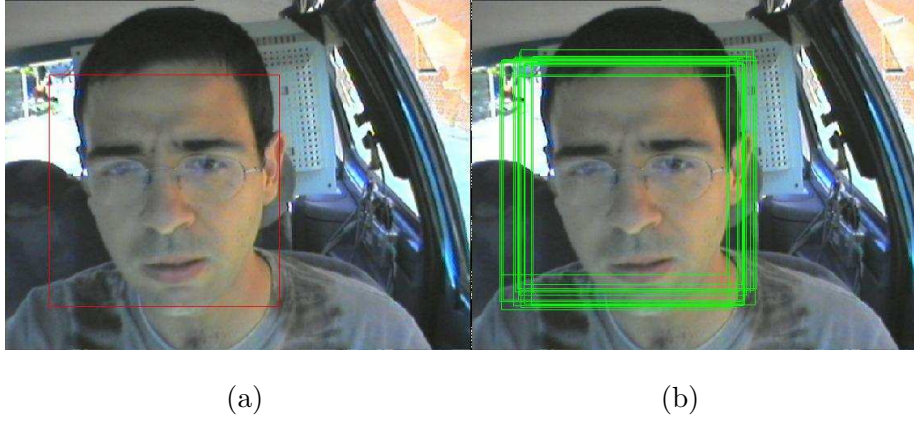


Figura 4.19: Representación de las partículas del algoritmo Condensation sobre el rostro del conductor, (a) región predicha, (b) partículas.

La evaluación de la densidad de las observaciones viene dada por la ecuación (4.21).

$$p(z_t|x_t) \propto \text{Cexp}(-\delta(R_1, R_2)) \quad (4.21)$$

donde $R_1 = R_1(x_{01}, y_{01}, w_1, h_1)$ es la región hipotética del rostro, $R_2 = R_2(x_{02}, y_{02}, w_2, h_2)$ es la región que contiene al rostro y que es localizada por la red de neuronas, y $\delta(\cdot, \cdot)$ se define por la ecuación (4.22).

$$\delta(R_1, R_2) = \frac{\Delta_x + \Delta_y}{\Delta_w + \Delta_h + 1} \quad (4.22)$$

donde $\Delta_x = |x_{01} - x_{02}|$, $\Delta_y = |y_{01} - y_{02}|$, $\Delta_w = \max(w_1, w_2)$ y $\Delta_h = \max(h_1, h_2)$.

4.5.5. Resultados experimentales del seguimiento del rostro

Para iniciar la ejecución del algoritmo Condensation es necesario conocer cierta información de antemano, en este caso, el número de partículas y la distribución inicial del estado, $p(x_0)$. En el primer caso se ha decidido utilizar $N = 30$ partículas, obtenidas experimentalmente. En el segundo caso, se utiliza la distribución normal de parámetros μ_0, Σ_0 , es decir, $p(x_0) \equiv N(\mu_0, \Sigma_0)$, donde el estado inicial puede ser descrito por (4.23) cuando se conoce el estado en dos imágenes consecutivas t y $t + 1$, y la matriz de varianza se establece por (4.24).

$$\begin{aligned}
x_0 &= x_{t+1} \\
y_0 &= y_{t+1} \\
u_0 &= x_{t+1} - x_t \\
v_0 &= y_{t+1} - y_t \\
w_0 &= w_{t+1} \\
h_0 &= h_{t+1}
\end{aligned} \tag{4.23}$$

$$\Sigma_0 = \begin{bmatrix} 100 & 0 & 0 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 & 0 & 0 \\ 0 & 0 & 25 & 0 & 0 & 0 \\ 0 & 0 & 0 & 25 & 0 & 0 \\ 0 & 0 & 0 & 0 & 9 & 0 \\ 0 & 0 & 0 & 0 & 0 & 9 \end{bmatrix} \tag{4.24}$$

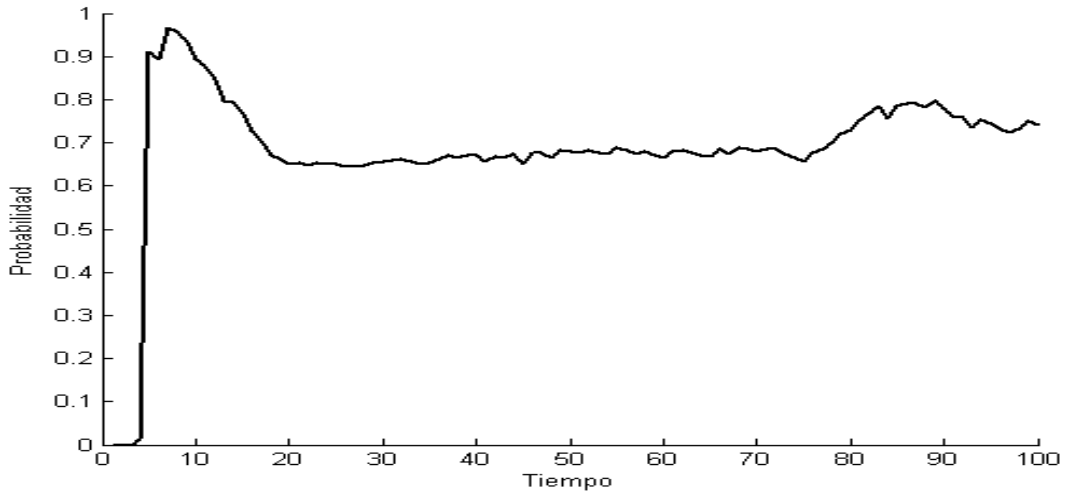


Figura 4.20: Probabilidad generada por Condensation en el seguimiento del rostro sobre una secuencia de 100 imágenes.

En la figura 4.20 se observan las probabilidades generadas con el algoritmo Condensation al seguir un rostro sobre una secuencia de 100 imágenes que genera la trayectoria de la figura 4.22 (a) y en la figura 4.22 (b) se observan los errores cometidos. A la vez, en la figura 4.21 se observan varios instantes del seguimiento del rostro bajo la metodología

4.5 Seguimiento del rostro



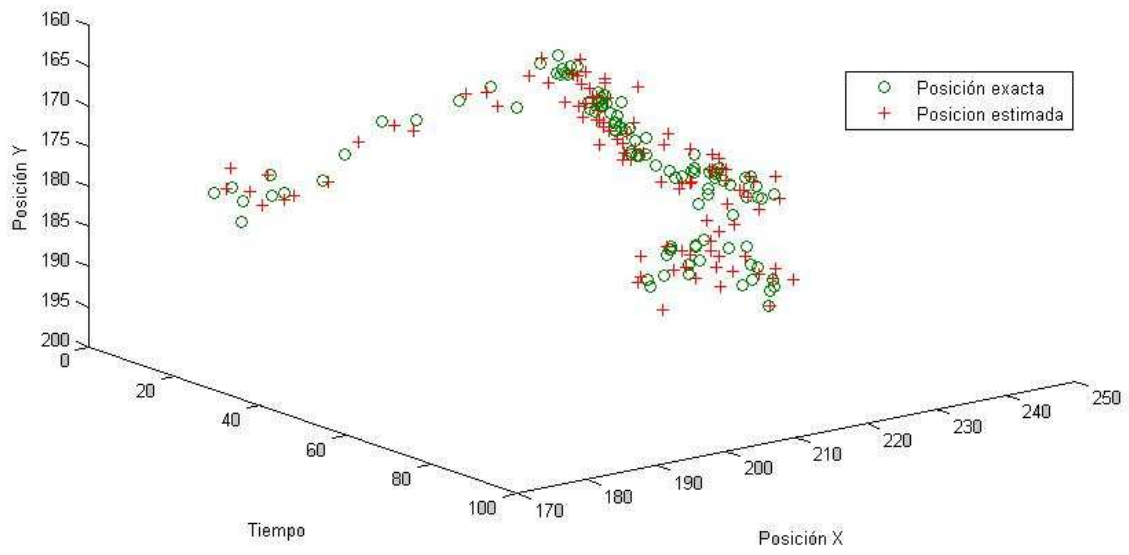
Figura 4.21: Corrección para detectar el rostro del conductor cuando el método de Viola & Jones falla.

propuesta.

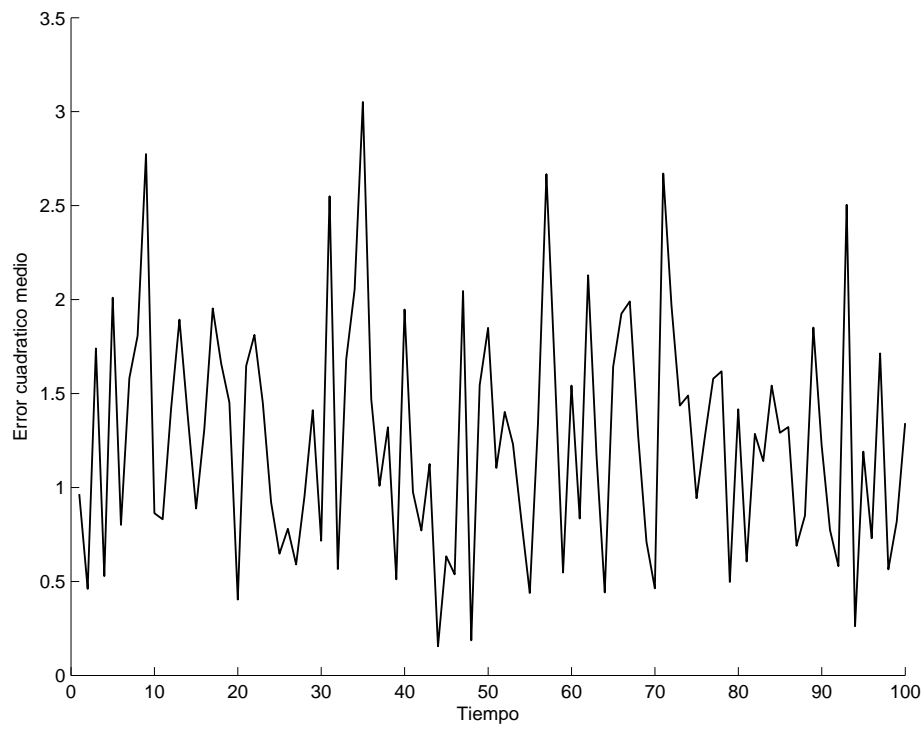
Para comprobar la validez del método propuesto se han desarrollado varios experimentos, estos incluyen varias secuencias de imágenes tomadas sobre el vehículo Ivvi, sus resultados se presentan en la tabla 4.1.

	Número de imágenes	Número de fallos	Aciertos (%)
Secuencia 1	960	60	93,75
Secuencia 2	900	22	97,55
Secuencia 3	330	15	95,45
Secuencia 4	500	45	91,00

Tabla 4.1: Resultados del seguimiento del rostro generado por el algoritmo Condensation sobre varias secuencias de imágenes de conductores sobre el vehículo Ivvi.



(a)



(b)

Figura 4.22: (a) Resultados y (b) error, generados con el algoritmo Condensation sobre una secuencia de 100 imágenes al estimar la posición del rostro.

4.6. Seguimiento de los ojos

Al igual que el caso anterior, el algoritmo de Condensation es implementado para el seguimiento de los ojos a través del tiempo. El seguimiento de los ojos constituye una parte delicada en el funcionamiento del sistema, porque para analizar el estado del conductor a través del análisis de los ojos es necesario mantenerlos siempre localizados.

En este caso, el seguimiento se realiza sobre los rectángulos RI_L y RI_R detectados en conjunto con el rostro.

4.6.1. Especificación del modelo de seguimiento de los ojos

En este caso el estado de cada ojo viene dado por la ecuación (4.25).

$$x_t = (x, y, u, v) \quad (4.25)$$

en esta ecuación (x, y) es la posición del centro del ojo y (u, v) es su velocidad en las direcciones x e y , respectivamente.

El modelo de transición que relaciona el estado presente con el estado inmediatamente anterior viene dado por la ecuación (4.26), que representa un proceso autoregresivo de orden uno $AR(1)$.

$$x_t = Ax_{t-1} + \xi_t \quad (4.26)$$

nuevamente, ξ_t representa la perturbación del sistema que viene dado por un ruido blanco y A es la matriz que genera la relación entre los dos instantes de tiempo, su estructura viene dada en [5], [42].

El modelo de medida viene dada por:

$$z_t = Hx_{t-1} + \eta_t \quad (4.27)$$

con η_t un ruido blanco, H viene dada en [3] y $z_t = (x, y)$ es la posición del centro el ojo.

4.6.2. Correlación de modelos

La correlación normalizada viene dada por la ecuación (4.28), cuyo valor es 1 para un emparejamiento perfecto y 0 en caso contrario, obviamente valores cercanos a 1 son considerados como satisfactorios.

$$g(x, y) = \frac{\sum ((I(x, y) - I_m)(M(x, y) - M_m))}{\sqrt{\sum (I(x, y) - I_m)^2 \sum (M(x, y) - M_m)^2}} \quad (4.28)$$

los valores I_m y M_m representan el valor promedio de las imágenes I y M , respectivamente.

Esto es lo que se denomina correlación de modelos (*template matching*), que consiste en calcular la similitud de una porción de imagen con respecto al modelo propuesto [28], [66], [131].

El modelo M (imagen utilizada de patrón) tiene dimensiones a y b para el ancho y el alto, respectivamente, por lo tanto, la suma de la ecuación (4.28) se realiza al amparo de estos dos parámetros sobre cada uno de los píxeles que corresponden a este dominio.

4.6.3. Evaluación de la función de densidad

En esta parte se ha decidido utilizar la técnica denominada *template matching*, porque en este caso el seguimiento se realiza sobre los rectángulos RI_R y RI_L generados en el módulo de detección del rostro, que se conoce que contienen un ojo. Por tal motivo, no es posible usar un método de clasificación similar al caso del rostro.

Para la evaluación de la función de densidad a posteriori se utilizó la correlación normalizada junto con el modelo generado con el método propuesto para detectar los ojos. El modelo que ingresa en la correlación es sometido a un preprosamiento a partir del cual se obtienen dos modelos, uno alargado verticalmente y otro alargado horizontalmente, ver figura 4.23.

Los pesos $\pi_t^{(n)}$ son aproximados por el valor de la correlación dada por la ecuación (4.28). La correlación es evaluada en la posición dada por cada una de las partículas generadas por el algoritmo Condensation, a través de la ecuación (4.29):

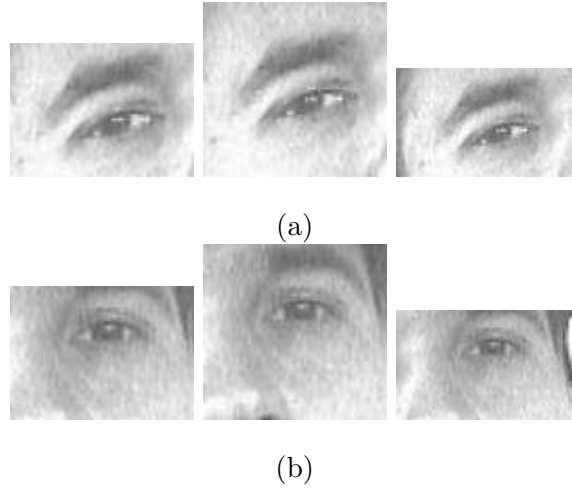


Figura 4.23: Modelos para la correlación de los ojos en el algoritmo Condensation (a) ojo derecho, (b) ojo izquierdo. Modelo inicial (columna izquierda), modelo alargado verticalmente (columna central) y modelo alargado horizontalmente (columna derecha).

$$p(z_t | x_t = s_t^{(n)}) = \max(g_1(s_t^{(n)}), g_2(s_t^{(n)}), g_3(s_t^{(n)})) \quad (4.29)$$

en este caso, $g_i(\cdot)$ para $i = 1, 2, 3$, representa el valor de la correlación del modelo i en la posición predicha por el algoritmo Condensation.

Al utilizar los modelos alargados se ha conseguido captar mejor las variaciones que sufren los ojos debido a los múltiples cambios que pueden presentarse, entre ellos, las variaciones por parpadear, los movimientos del rostro, etc.

4.6.4. Resultados experimentales del seguimiento de los ojos

Para iniciar la ejecución del algoritmo Condensation se debe conocer la distribución inicial del estado, es decir, conocer $p(x_0)$; en este caso, se utiliza la distribución normal de parámetros μ_0, Σ_0 , es decir, $p(x_0) \equiv N(\mu_0, \Sigma_0)$, donde el estado inicial puede ser descrito por (4.30) cuando se conoce el estado en dos imágenes consecutivas t y $t + 1$, y la matriz de varianza se establece por (4.31). El número de partículas utilizado es de $N = 50$, con lo cual se obtienen resultados satisfactorios tanto en fiabilidad como en tiempo de cómputo.

$$\begin{aligned}
x_0 &= x_{t+1} \\
y_0 &= y_{t+1} \\
u_0 &= x_{t+1} - x_t \\
v_0 &= y_{t+1} - y_t
\end{aligned} \tag{4.30}$$

$$\Sigma_0 = \begin{bmatrix} 100 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 \\ 0 & 0 & 25 & 0 \\ 0 & 0 & 0 & 25 \end{bmatrix} \tag{4.31}$$

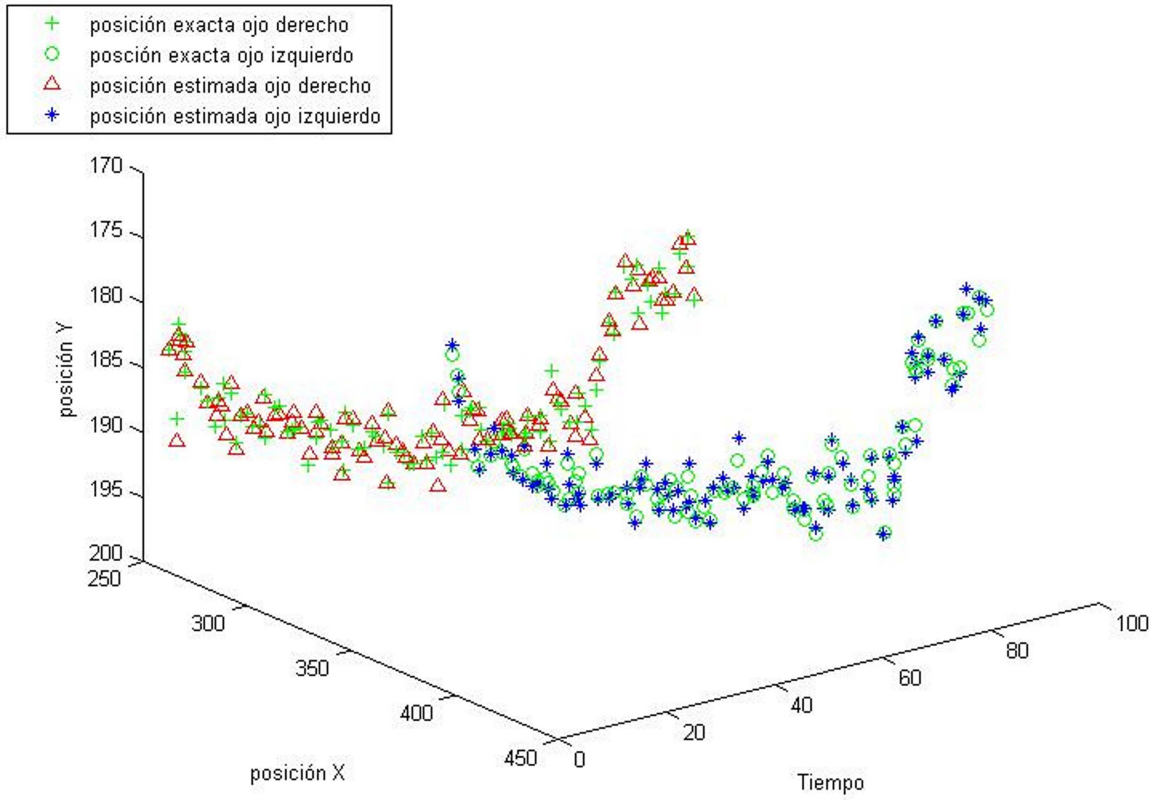


Figura 4.24: Resultados del algoritmo Condensation sobre una secuencia de 100 imágenes para estimar la posición de los ojos.

En la figura 4.24 se presentan los resultados del seguimiento desarrollado sobre una secuencia de 100 imágenes consecutivas tomadas a un conductor y en la figura 4.26 se

4.6 Seguimiento de los ojos

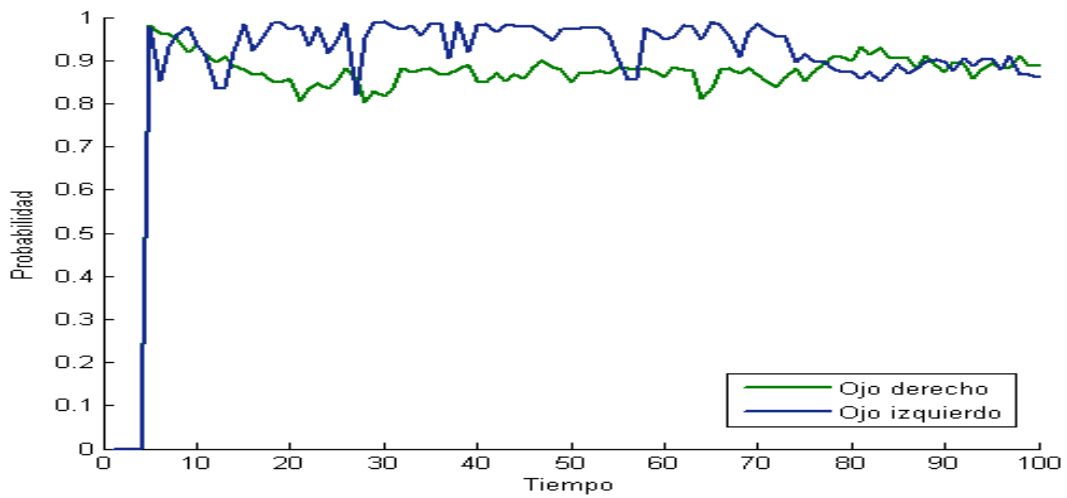


Figura 4.25: Probabilidad generada por Condensation en el seguimiento de los ojos derecho e izquierdo.

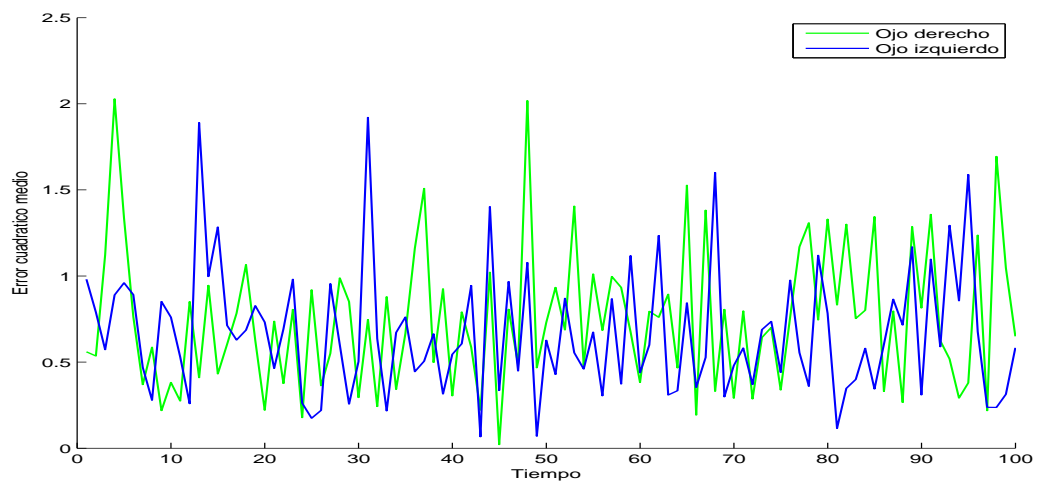


Figura 4.26: Error generado por Condensation en el seguimiento de los ojos derecho e izquierdo.

muestran sus errores (error cuadrático medio). Cabe anotar que se implementa un algoritmo Condensation para el ojo izquierdo y uno para el ojo derecho.

Para finalizar esta sección, en la tabla 4.2 se aprecian los resultados obtenidos con el proceso de seguimiento de los ojos sobre cuatro secuencias compuestas de varias imágenes



Figura 4.27: Resultados del seguimiento de los ojos del conductor bajo el algoritmo Condensation.

de conductores en condiciones reales de conducción, y en la figura 4.27 se observa varios instantes sobre una secuencia de 100 fotogramas.

	Número de imágenes	Número de fallos	Aciertos (%)
Secuencia 1	960	20	97,91
Secuencia 2	900	30	96,60
Secuencia 3	330	14	95,75
Secuencia 4	500	8	98,40

Tabla 4.2: Resultados del seguimiento de los ojos generado por el algoritmo Condensation sobre varias secuencias de imágenes de conductores sobre el vehículo IvvI.

4.7. Cálculo del índice de somnolencia con SVM

PERCLOS es el índice más utilizado en la construcción de este tipo de sistemas, por tal motivo ha sido implementado en este trabajo. Sin embargo, existen otros índices como el AECS [124] y la duración de cierre del ojo durante cinco fotogramas consecutivos [152], que también son utilizados para éste propósito.

Para calcular los índices es imprescindible conocer el estado del ojo en cada instante.

4.7 Cálculo del índice de somnolencia con SVM

En este trabajo se manejan dos estados: abierto y cerrado. Para conocer estos estados se utilizará una técnica robusta de clasificación denominada SVM, acrónimo del inglés Support Vector Machine que en español suele tener el nombre de Máquina de Vectores Soporte, en conjunto con el Filtro de Gabor, técnica ampliamente utilizada para contribuir en el realce de ciertas características.

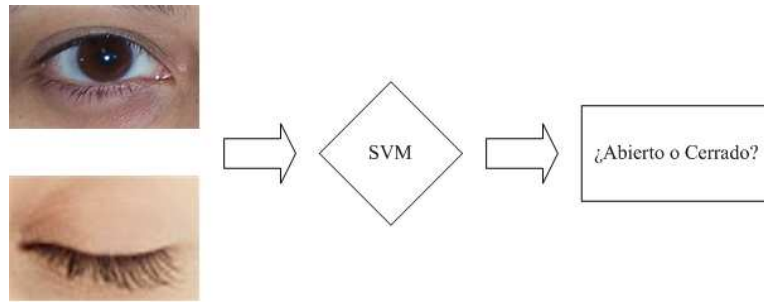


Figura 4.28: Esquema de consulta a SVM.

En la figura 4.28 se ilustra el proceso de clasificación entre ojo abierto y ojo cerrado utilizando SVM. De acuerdo a esta figura se nota que SVM se convierte en una “caja negra” que procesa la información para entregar un resultado sin conocer cual es el procedimiento empleado, por tal motivo, es importante resaltar las características que mejor identifican a los objetos que se desea estudiar.

4.7.1. Descripción de SVM

Partiendo de los estados que se precisa conocer se propone utilizar un método de aprendizaje supervisado [106], [64], [11] en este caso, support vector machine (*SVM*); que fue introducido por Boser *et al.* [41] en 1.992 y desde esa fecha ha sido ampliamente utilizado en varias aplicaciones, entre ellas, para resolver problemas de agrupamiento (clustering), selección de características y regresión [64].

Matemáticamente *SVM* se plantea a través del siguiente problema de optimización cuadrática.

$$\begin{aligned}
\min_{\alpha} J(\alpha) &= \alpha Q \alpha^T - e^T \alpha \\
s.t. \\
0 &\leq \alpha_i \leq C \text{ para } i = 1, \dots, m \\
y^T \alpha &= 0
\end{aligned} \tag{4.32}$$

donde $e \in \mathbb{R}^m$ es el vector de unos, C es una cota superior que penaliza el error, Q es una matriz de tamaño $m \times m$ con $Q_{ij} = y_i y_j K(x_i, x_j)$, $K(x_i, x_j)$ es la denominada función núcleo, que puede ser una función lineal o no-lineal, x_i es el vector de características y y es el vector que representa las clases.

Al resolver el problema de optimización planteado en la ecuación (4.32), *SVM* se reduce a encontrar el hiperplano que mejor discrimina entre las clases involucradas y a la vez reduce el riesgo empírico [106], [76]. Este proceso se encuentra representado en la figura 4.29.

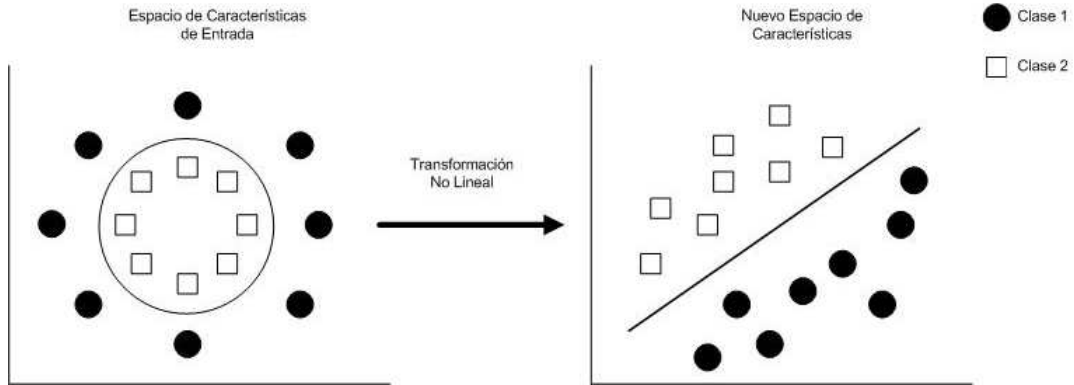


Figura 4.29: Representación de la optimización desarrollada en SVM.

4.7.2. Filtro de Gabor

Fueron introducidos por Daugman [54] en el análisis de imágenes para resaltar ciertas características al variar la orientación y la escala. El filtro de Gabor se forma a partir de los wavelets Gabor, también conocidos como núcleos o filtros [20], [101]. La función que genera el wavelet Gabor viene definida por:

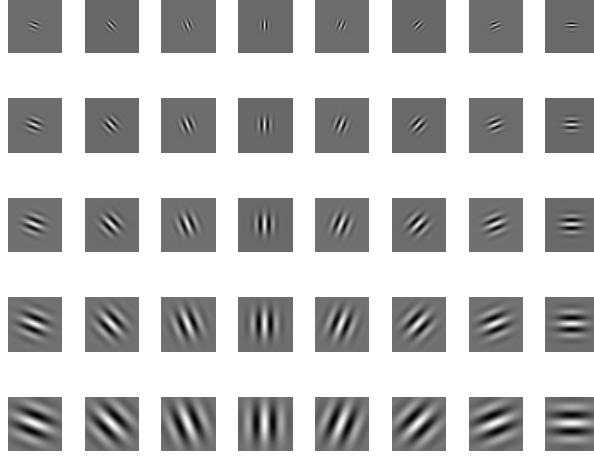


Figura 4.30: Representación del wavelet de Gabor con 5 escalas (filas) y 8 orientaciones (columnas).

$$\Psi_{\mu,\nu}(z) = \frac{\|k_{\mu,\nu}\|^2}{\sigma^2} \exp\left(-\frac{\|k_{\mu,\nu}\|^2 \|z\|^2}{2\sigma^2}\right) \left[\exp(ik_{\mu,\nu}z) - \exp\left(-\frac{\sigma^2}{2}\right) \right] \quad (4.33)$$

donde $z = (x, y)$ es el punto de coordenadas x e y . Los parámetros μ y ν definen la orientación y la escala, $\|\cdot\|$ define la norma, y σ es la desviación estandar de la envolvente gaussiana. El vector complejo que representa la onda $k_{\mu,\nu}$ se define por:

$$k_{\mu,\nu} = k_{\nu} \exp(i\phi_{\mu}) \quad (4.34)$$

donde $i = \sqrt{-1}$ es el número complejo, $k_{\nu} = \frac{k_{max}}{f^{\nu}}$, $\phi_{\mu} = \frac{\pi\mu}{8}$, k_{max} es la frecuencia máxima y f es el espacio entre los núcleos en el dominio de la frecuencia.

Los valores de escala y orientación comúnmente utilizados en varias aplicaciones en distintos campos son $\nu = \{0, 1, 2, 3, 4\}$ y $\mu = \{0, 1, \dots, 7\}$ [20]; cuya representación se puede observar en la figura 4.30.

4.7.3. Representación 2D del ojo a través del filtro de Gabor

La representación de una imagen bajo el wavelet de Gabor es la convolución de la imagen con la familia de núcleos Gabor. La convolución de la imagen $I(z)$ y el wavelet de Gabor $\Psi_{\mu,\nu}(z)$ se definida a través de la integral de la ecuación (4.35).

$$G_{\mu,\nu}(z) = I(z) * \Psi_{\mu,\nu}(z) = \int I(t) \Psi_{\mu,\nu}(z - t) dt \quad (4.35)$$

Al variar los parámetros de orientación y escala se obtiene el conjunto de representaciones Gabor de la imagen $I(z)$ que se denota por $\mathcal{S} = \{G_{\mu,\nu}(z) : \mu = 0, 1, \dots, n_\mu, \nu = 0, 1, \dots, n_\nu, z = (x, y)\}$.

La respuesta $G_{\mu,\nu}(z)$ de cada núcleo Gabor es una función compleja con una parte real y una parte imaginaria cuya magnitud viene dada por:

$$\|G_{\mu,\nu}(z)\| = \sqrt{\text{real}(G_{\mu,\nu}(z))^2 + \text{imag}(G_{\mu,\nu}(z))^2} \quad (4.36)$$

La magnitud de las imágenes generadas con el wavelet Gabor es utilizada para construir las características que representan los estados del ojo. De aquí en adelante, $\|G_{\mu,\nu}(z)\|$ se denominará magnitud Gabor.

En este caso particular se trata de resaltar las características que mejor representan tanto al ojo abierto como al ojo cerrado. Tomando en cuenta los inconvenientes descritos anteriormente, se añaden los que se generan por la posición de la cámara, que introduce problemas de distancia y distorsión.

Experimentalmente se ha logrado comprobar que a ciertas escalas y orientaciones del filtro de Gabor, la zona blanca del ojo (esclerosis) y los bordes de los párpados sobresalen dentro de la imagen. Por esta razón se ha utilizado tres escalas y cuatro orientaciones, ellas son $\{\pi/8, \pi/4, \pi/2\}$ para la escala y $\{0, \pi/6, 2\pi/6, 3\pi/6\}$ para la orientación. Esto produce 12 imágenes de salida de las cuales se extraen las componentes que forman el vector de características. La figura 4.31 muestra el resultado de aplicar el filtro sobre la imagen de un ojo abierto y un ojo cerrado.

4.7.4. Extracción de características basada en el filtro de Gabor

La extracción de características se realiza sobre las imágenes de la magnitud del conjunto \mathcal{S} generadas con el filtro de Gabor. Con los elementos de \mathcal{S} se construye el vector de características que es la entrada del clasificador diseñado para identificar el estado de los ojos del conductor.

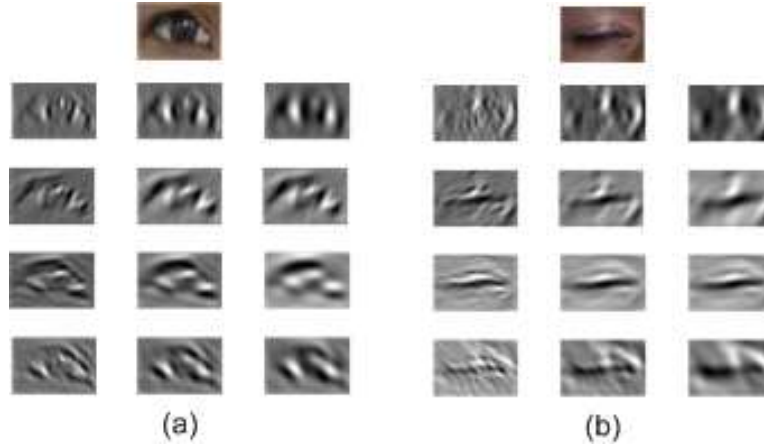


Figura 4.31: Imágenes generadas con el filtro Gabor con tres escalas y cuatro orientaciones, (a) ojo abierto, (b) ojo cerrado.

Las imágenes que forman la base de datos de los ojos tienen dimensiones 30 por 20 píxeles para el ancho y el alto, respectivamente. Para extraer las características se divide cada una las imágenes Gabor en 20 sub-imágenes (ventanas) de tamaño 6 por 5, a partir de aquí, el vector de características se forma con el valor promedio calculado sobre cada una de las celdas. El proceso de división en sub-ventanas se ve gráficamente en la figura 4.32. Este procedimiento tiene relación con varias investigaciones anteriores y es una variación de los trabajos de Chen y Kubo [151], Ray *et al.* [98], Gökberk *et al.* [9] y es una extensión del filtro presentado por Feng y Yuen [13] denominado *eye variance filter*.

Bajo esta afirmación, sea D el vector de características en \mathbb{R}^{360} con coordenadas $d_i^{\mu,\nu}$ donde:

$$d_i^{\mu,\nu} = \frac{1}{30} \sum_{y=1}^5 \sum_{x=1}^6 \|G_{\mu,\nu}(x,y)\| \text{ para } i = 1, \dots, 20 \quad (4.37)$$

los valores ν y μ corresponden a los parámetros de orientación y escala del filtro de Gabor.

4.7.5. Entrenamiento de SVM

Para realizar este trabajo se ha construido una base de datos que consta de ojos abiertos y ojos cerrados, ejemplos que forman parte de ésta base se pueden apreciar en la figura 4.33. Las imágenes provienen de varias fuentes (conductores sobre el vehículo

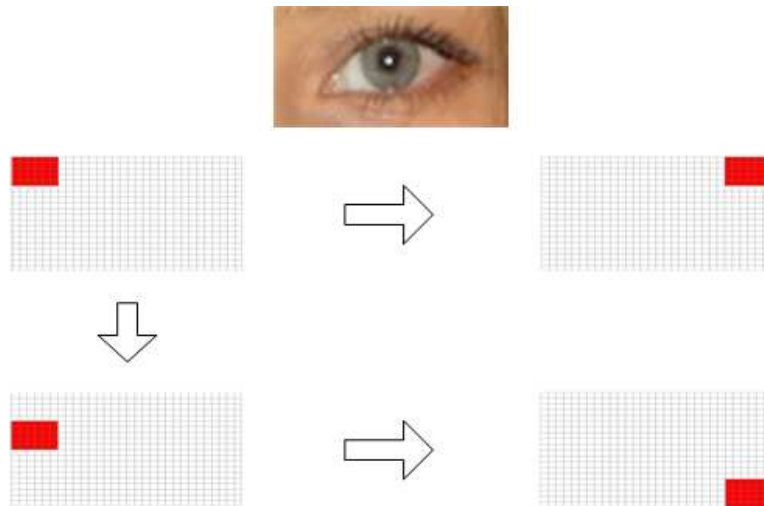


Figura 4.32: Ventanas para extraer el vector de características de la imagen del ojo.

IvvI, internet, fotos, etc.) con distintas condiciones de iluminación y diferentes razas, esto permite cubrir la mayor variedad de tipos de color de ojos existentes.

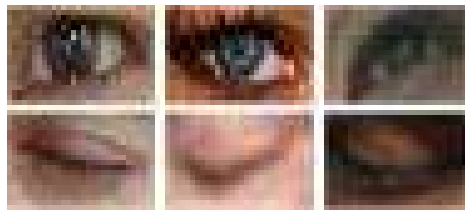


Figura 4.33: Ejemplos de la base de datos de ojos.

Previo al entrenamiento de *SVM* es indispensable procesar cada imagen para resaltar, en la medida de lo posible, las características que identifican a un ojo abierto de un ojo cerrado. El preprocesamiento consiste en los siguientes pasos:

1. Equalizar la imagen para homogenizar y reducir los cambios de iluminación.
2. Filtra la imagen con el filtro de la mediana para reducir el ruido.
3. Filtrar la imagen con el filtro Sharpen para resaltar los bordes.
4. Aplicar el filtro de Gabor.
5. Normalizar la imagen al intervalo $[0, 1]$.

4.7 Cálculo del índice de somnolencia con SVM

El principal objetivo de entrenar *SVM* es encontrar los mejores parámetros y el mejor núcleo que permitan minimizar la ecuación (4.32). Cabe mencionar que existen varias funciones utilizadas habitualmente como núcleo, de entre ellas se ha escogido la función *radial basic function* (RBF), que es una función no lineal, en la figura 4.34 se observa éste núcleo en el caso de dos dimensiones.

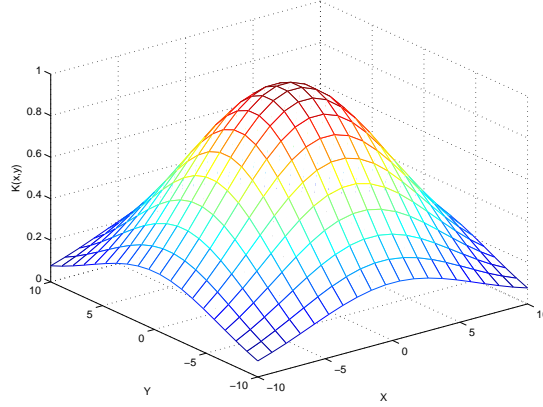


Figura 4.34: Función RBF en dos dimensiones.

Una vez escogido el núcleo y luego de desarrollar varios experimentos se obtienen los siguientes parámetros:

- $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$,
- $C = 35$, y
- $\gamma = 0,0128$

donde $\exp(\cdot)$ es la función exponencial, y $\|\cdot\|$ es la norma euclídea.

Estos parámetros permiten alcanzar una alta tasa de clasificación que está alrededor del 94 % sobre una base de datos de pruebas que consta de 500 ojos y 1000 no-ojos.

En las figuras que se presentan a continuación se pretende verificar y exponer la potencia de este índice en dos casos:

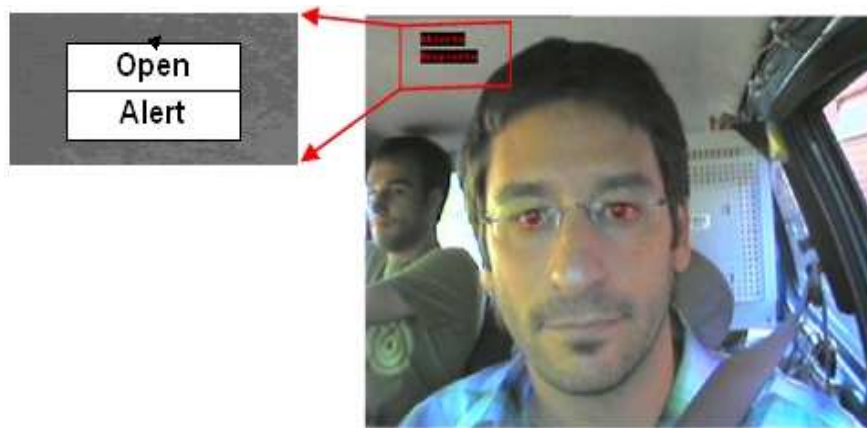


Figura 4.35: Resultado del análisis instantáneo del estado del conductor.

- conductor somnoliento, y
- conductor en pleno estado de vigilia

Para validar la eficacia de la clasificación realizada por *SVM*, en la tabla 4.3 se presentan los resultados obtenidos al aplicar el algoritmo sobre cuatro secuencias de imágenes. Los resultados son bastante satisfactorios debido a que se alcanza un alto porcentaje de aciertos en la clasificación del estado del ojo.

	Número de imágenes	Ojos Abiertos	Ojos Cerrados	Aciertos (%)
Secuencia 1	960	690/700	258/260	98,90
Secuencia 2	900	520/560	339/340	96,28
Secuencia 3	330	150/170	152/160	91,61
Secuencia 4	500	388/400	99/100	98,00

Tabla 4.3: Resultados del análisis del estado de los ojos usando *SVM* sobre varias secuencias de imágenes de conductores.

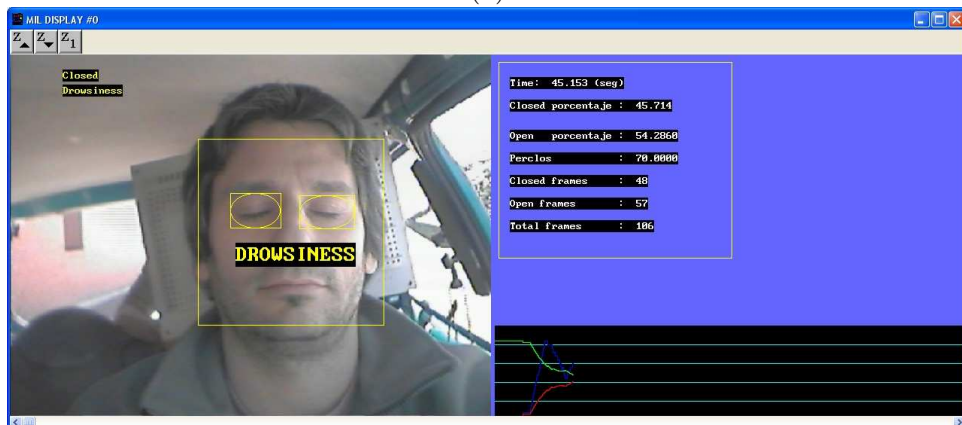
4.7 Cálculo del índice de somnolencia con SVM

4.7.6. Construcción del índice de somnolencia

Una vez que se puede clasificar el estado de los ojos entre abierto y cerrado, se construyen los índices de somnolencia propuestos por Horng et. al. [153] y Dong and Wu [152], figura 4.37, y PERCLOS, figura 4.38. El primero implican que si durante cinco fotogramas consecutivos se detecta que los ojos están cerrados se emite una alarma de alerta ante un posible estado de adormecimiento. El segundo, al sobrepasar un determinado umbral detecta síntomas de somnolencia.



(a)



(b)

Figura 4.36: Resultado del análisis del estado del conductor sobre el tiempo, (a) estado de alerta, (b) estado de somnolencia.

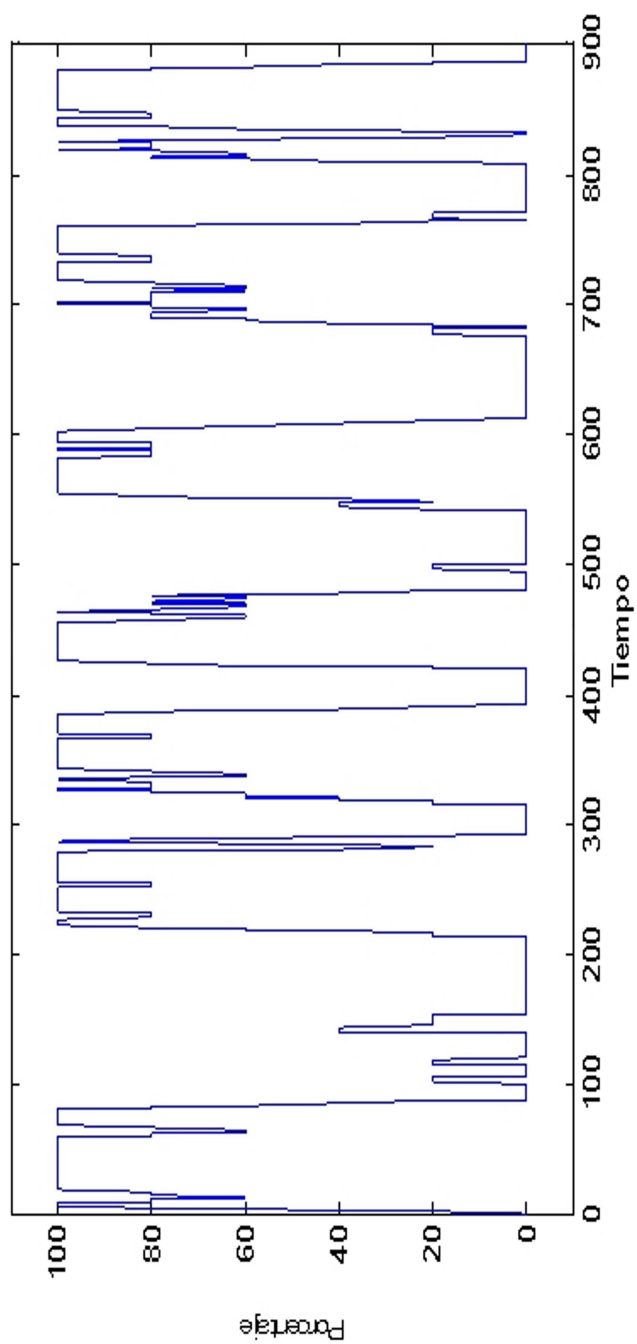


Figura 4.37: Evolución del índice de Horng *et al.* y Dong-Wu en una secuencia de 900 imágenes.

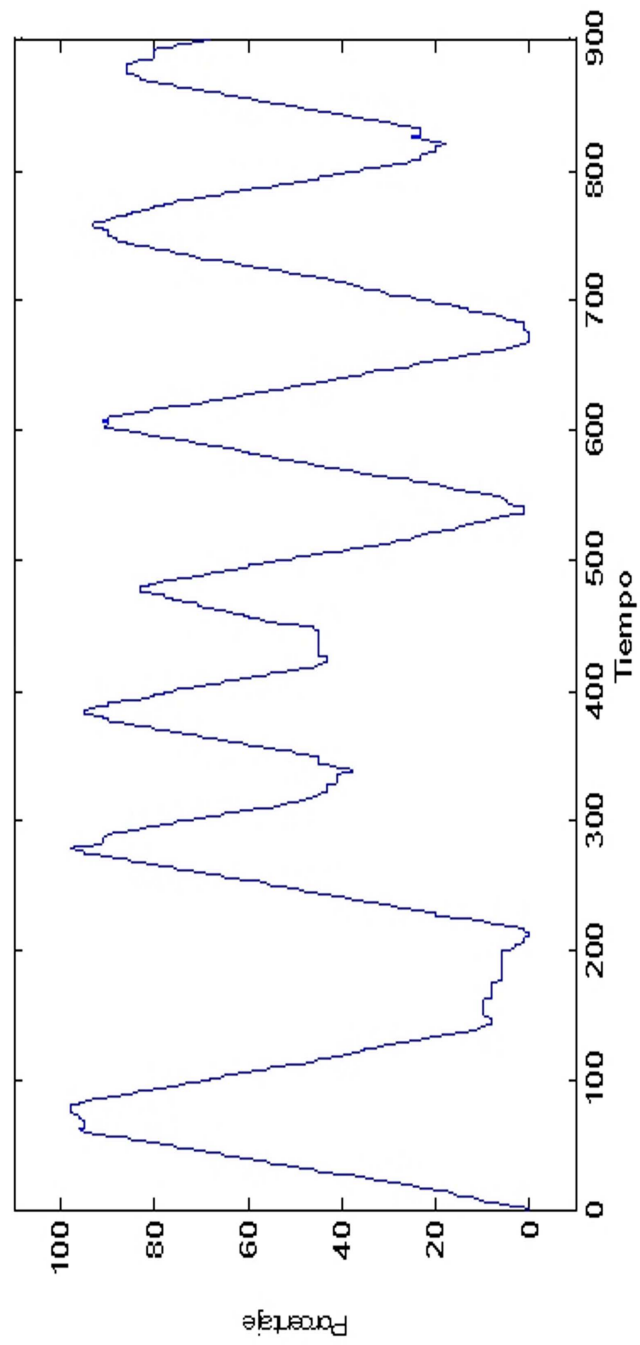


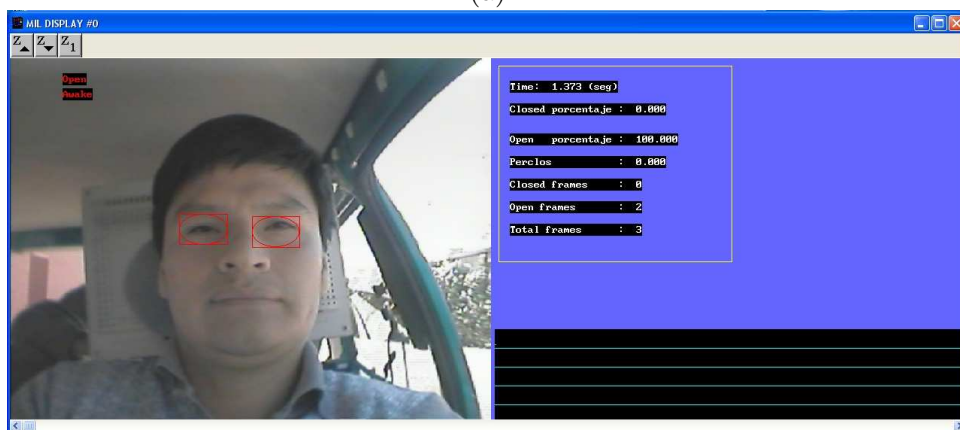
Figura 4.38: Evolución de Perclos en una secuencia de 900 imágenes.

Al analizar las figuras 4.37 y 4.38 se aprecian varios picos que se forman sobre los índices calculados a partir de un conductor que presenta alto grado de somnolencia. En ambos casos, el sistema es capaz de detectar presencia de adormecimiento, figuras 4.36 (b) y 4.39 (a).

Adicionalmente, en el diseño final del sistema, que se presenta en la figura 4.36 y en la figura 4.39 se observan tres gráficos de evolución en el panel inferior derecho. El primero corresponde al índice de somnolencia (azul); el segundo, calcula el porcentaje de veces que los ojos se encuentran abiertos (verde); y el tercero, mide el porcentaje de veces que los ojos se encuentran cerrados (rojo).



(a)



(b)

Figura 4.39: Resultado del análisis del estado del conductor sobre el tiempo, (a) estado de somnolencia, (b) estado de alerta.

4.8. Detección de la distracción

4.8.1. Orientación del rostro

Para estimar la orientación del rostro a partir de la posición de los ojos se utiliza la ecuación (4.38), que mide el ángulo que forma el rostro respecto al eje vertical, la figura 4.40 muestra gráficamente lo descrito.

$$\theta = \tan^{-1} \left(\frac{y_2 - y_1}{x_2 - x_1} \right) \quad (4.38)$$

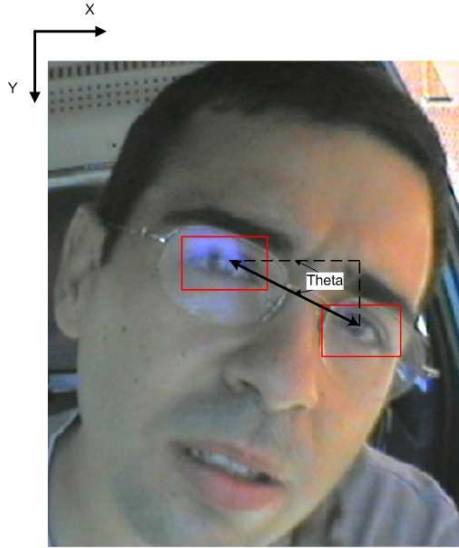


Figura 4.40: Estimación de la orientación del rostro a partir de la posición de los ojos.

Experimentalmente se ha comprobado que los límites fijados por la ecuación (4.39) dan una buena estimación para conocer la orientación del rostro.

$$\begin{cases} \text{Izquierda} & \text{si } \theta > 8 \\ \text{Frente} & \text{si } |\theta| \leq 8 \\ \text{Derecha} & \text{si } \theta < -8 \end{cases} \quad (4.39)$$

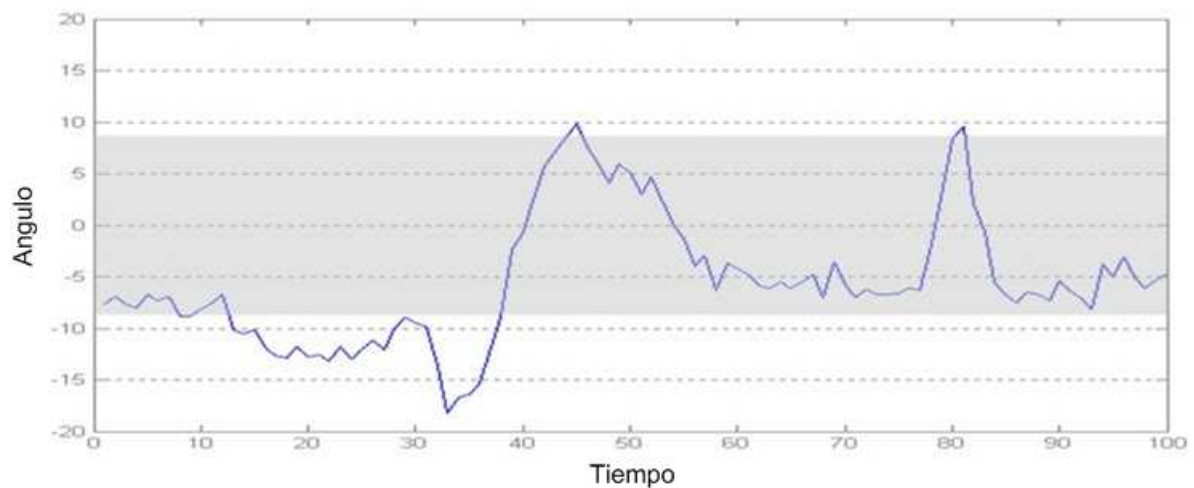


Figura 4.41: Evolución de la orientación del rostro a partir de la posición de los ojos.

4.8.2. Dirección del rostro usando redes neuronales

El análisis de la distracción es importante porque ella ha generado muchos accidentes y ha contribuido de manera negativa en el aumento de las estadísticas de accidentes de tráfico, tal como se presentó en el capítulo donde se presentaron varias estadísticas que reflejan este percance.

El índice de distracción se calcula una vez que el sistema detecta que el rostro no está de frente a la cámara durante 5 fotogramas consecutivos, por tal motivo, el seguimiento del rostro es útil para detectar si el conductor se encuentra distraído.

Para comprobar que el rostro no está de frente a la cámara se utiliza la red neuronal que fue descrita en la sección anterior, un ejemplo de la respuesta del sistema se observa en la figura 4.42, donde se observa que una flecha indica la dirección del rostro, categorizada en cinco estados: derecha, media-derecha, frente, media-izquierda e izquierda.

4.9 Velocidad de procesamiento



Figura 4.42: Resultados del análisis de la distracción del conductor sobre el tiempo, (a) medio giro a derecha, (b) giro a derecha, (c) medio giro a izquierda, (d) giro a izquierda.

4.9. Velocidad de procesamiento

En los experimentos realizados, la velocidad promedio de procesamiento fue de 325 milisegundos por fotograma, es decir, entre 3 y 4 fotogramas por segundo. Esto incluye la captura del fotograma, el procesamiento y extracción de la información y, finalmente, la visualización de los resultados por pantalla. Estos experimentos se desarrollaron en un Intel Pentium D, con 3,2 GHz. de velocidad, 2 GB. de memoria RAM y MS Windows XP.

En la tabla 4.4 se incluye el promedio de tiempo de cómputo de las pruebas realizadas.

Etapa	Tiempo (ms.)
Detección del rostro	350
Detección de los ojos	40
Seguimiento del rostro	130
Seguimiento de los ojos	120
Análisis del estado de los ojos	15
Análisis distracción	15
Presentación de resultados	5

Tabla 4.4: Tiempos de cómputo en cada fase del algoritmo en el espectro visible.

Detección de la Somnolencia en el Infrarrojo Cercano

Para que un sistema de detección de la somnolencia pueda ser embarcado en un vehículo real debe funcionar en todo instante, por este motivo, para trabajar en la noche se propone un sistema que se compone de un sistema de visión basado en iluminación infrarroja y de varios algoritmos para aprovechar el efecto que esta luz produce sobre las pupilas. El sistema de visión fue presentado en el capítulo tres; mientras que el algoritmo será desarrollado ampliamente en el presente. Con esto se pretende dar una solución al problema de la somnolencia del conductor por la noche, y a la vez, complementar el sistema del día.

5.1. Descripción del sistema en el infrarrojo cercano: diseño y alcance

A partir de la información presentada en la introducción se concluye que gran parte de los accidentes de tráfico producidos por la somnolencia se presentan en la noche, hasta el punto que se puede asegurar que la mayor frecuencia de accidentes se localizan en la franja horaria que fluctúa entre la media noche y el amanecer. Este alto índice de siniestros

comprendidos en este intervalo de tiempo ha motivado la construcción de un sistema que sea capaz de operar en la noche, donde las nuevas condiciones de iluminación imponen nuevos retos a ser superados durante esta parte de la investigación.

Para localizar los ojos en la noche se utiliza en esencia la iluminación infrarroja, en conjunto con las propiedades de reflexión de dicha luz sobre las pupilas de los ojos. La luz infrarroja al incidir sobre la pupila la resalta, permitiendo localizar su posición y facilitando su posterior análisis, como se verá más adelante a lo largo de este capítulo.

En la figura 5.1 se presenta el esquema del sistema de detección de la somnolencia en la noche, antes de seguir es necesario hacer énfasis en las características de las imágenes de entrada, ellas se encuentran saturadas por el efecto de la iluminación infrarroja, lo que genera imágenes monocromáticas de baja calidad comparadas con las imágenes que se podrían obtener en el espectro visible [162].

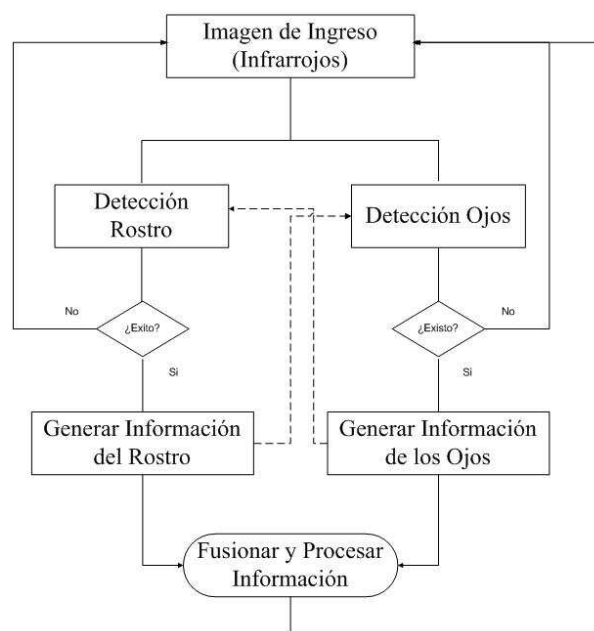


Figura 5.1: Esquema del algoritmo de detección de la somnolencia en el infrarrojo cercano.

Siguiendo la misma metodología del caso en el espectro visible, este sistema se compone de cinco módulos, en el siguiente orden:

- Detección de los ojos

5.1 Descripción del sistema en el infrarrojo cercano: diseño y alcance

- Detección del rostro
- Seguimiento de los ojos y el rostro
- Detección de la somnolencia
- Detección de la distracción

Para desarrollar esta parte de la investigación se utilizan las imágenes capturas con el sistema de visión descrito en el capítulo tres; este sistema de percepción utiliza iluminación infrarroja y captura imágenes similares a las presentadas en la figura 5.2.



Figura 5.2: Imágenes de distintos conductores generadas con el sistema de visión bajo iluminación infrarroja instalado sobre el vehículo IvvI.

Por otra parte, para apreciar mejor el fenómeno físico generado por el sistema de percepción, en la imagen de la figura 5.3 se muestra una sección ampliada en la que se ve como alternan verticalmente las líneas de los campos par e impar sobre la zona del ojo.

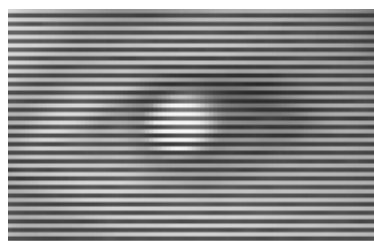


Figura 5.3: Estructura de la imagen capturada con el sistema de visión con iluminación infrarroja.

Una parte esencial del sistema consiste en separar los campos par e impar de la imagen tomada con infrarrojos y que constituye la base para la detección de los ojos, por este motivo, en este capítulo se comienza por describir el método de detección de los ojos para luego seguir con el método de detección del rostro, justo al contrario que en el caso del espectro visible.

5.2. Detección de los ojos

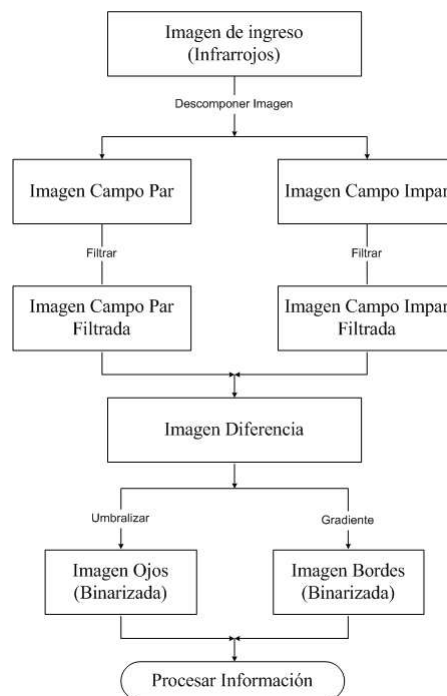


Figura 5.4: Diagrama para detectar los ojos en el infrarrojo cercano.

Para detectar los ojos sobre las imágenes generadas con iluminación infrarroja se procede de acuerdo al diagrama de la figura 5.4 que consiste esencialmente de tres partes:

- Generar la imagen diferencia y su imagen de bordes
- Construir un método de umbralización para la imagen diferencia, y para la imagen de bordes
- Localizar la posición de los ojos

5.2 Detección de los ojos

5.2.1. La imagen diferencia

En el diagrama presentado en la figura 5.4 se observa que la imagen de entrada se descompone en dos imágenes, denominadas imagen del campo par e imagen del campo impar (ver figura 5.5), cuyos nombres hacen referencia a la información del circuito integrado LM1881N. Este paso se realiza a través de las ecuaciones (5.1) y (5.2). Sea $I(x, y)$ el valor de una imagen generada con el sistema de iluminación infrarroja en el punto (x, y) .

$$I_P(x, y) = I(x, 2y - 1) \quad y = 1, \dots, h/2 \quad (5.1)$$

$$I_I(x, y) = I(x, 2y) \quad y = 1, \dots, h/2 \quad (5.2)$$

donde I_P e I_I denotan las imágenes del campo par e impar en el punto (x, y) , respectivamente, h es el alto de la imagen original. Las imágenes I_P e I_I tienen la mitad del alto de la imagen original I .

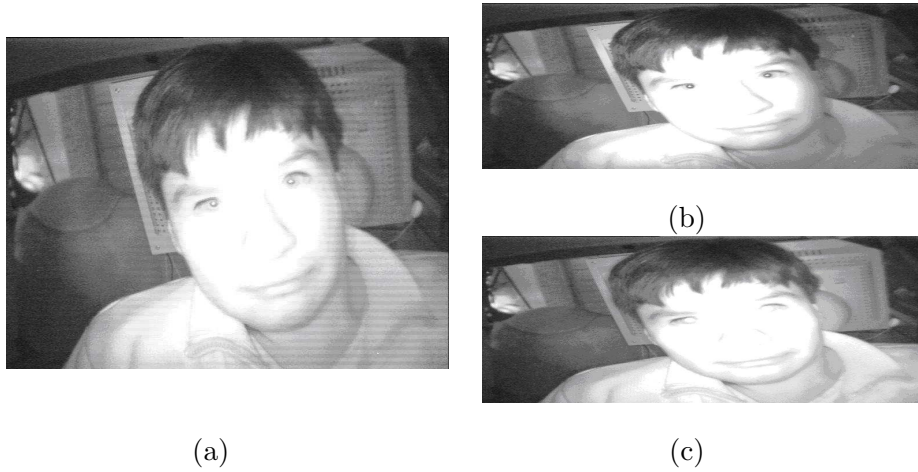


Figura 5.5: Separación de la imagen en el campo par e impar, (a) imagen capturada con el sistema de visión con iluminación infrarroja, (b) imagen del campo par, (c) imagen del campo impar.

Para filtrar las imágenes se utiliza el filtro de la media, representado por la máscara de la ecuación (5.3).

$$M = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \quad (5.3)$$

El filtro de la media es un filtro lineal espacial que genera un valor promedio calculado sobre la zona de la ventana de influencia. Con este proceso se logra eliminar el ruido que corresponde a un suavizado o filtrado espacial paso bajo que borra los detalles más finos de la imagen que se comportan como ruido de altas frecuencias [28].

Con las dos imágenes preparadas, es decir, filtradas en el paso anterior, se procede a construir la *imagen diferencia*, I_D , usando la ecuación (5.4), que representa el valor absoluto de la diferencia de cada uno de los píxeles de las imágenes involucradas. En la figura 5.6 (a-d-g) se muestran varios ejemplos de lo expuesto.

$$I_D(x, y) = |I_P(x, y) - I_I(x, y)| \quad (5.4)$$

Continuando con el desarrollo, al analizar la información de los histogramas que generan estas imágenes se verifica que es posible obtener un umbral que permite separar los objetos de interés (los ojos) del resto (fondo). En principio se puede usar un umbral fijo para realizar este proceso; pero debido a las condiciones cambiantes del ambiente se ha utilizado el algoritmo que se propone a continuación.

5.2.2. Método propuesto para umbralizar de la imagen diferencia

Para umbralizar la imagen diferencia, es decir, para separar las zonas de mayor valor de las zonas de menor valor, en la escala de grises, se propone el siguiente algoritmo:

1. Eliminar los píxeles del intervalo entre 0 y 20 para evitar distorsiones en el histograma porque son de interés los píxeles cercanos al color blanco
2. Seleccionar un umbral inicial T , en este caso tomar el valor de 127
3. Generar dos grupos de píxeles $G1$ y $G2$ con la siguiente restricción: $G1$ corresponde a los píxeles mayores a T , mientras que $G2$ corresponde a los píxeles menores o iguales a T

5.2 Detección de los ojos

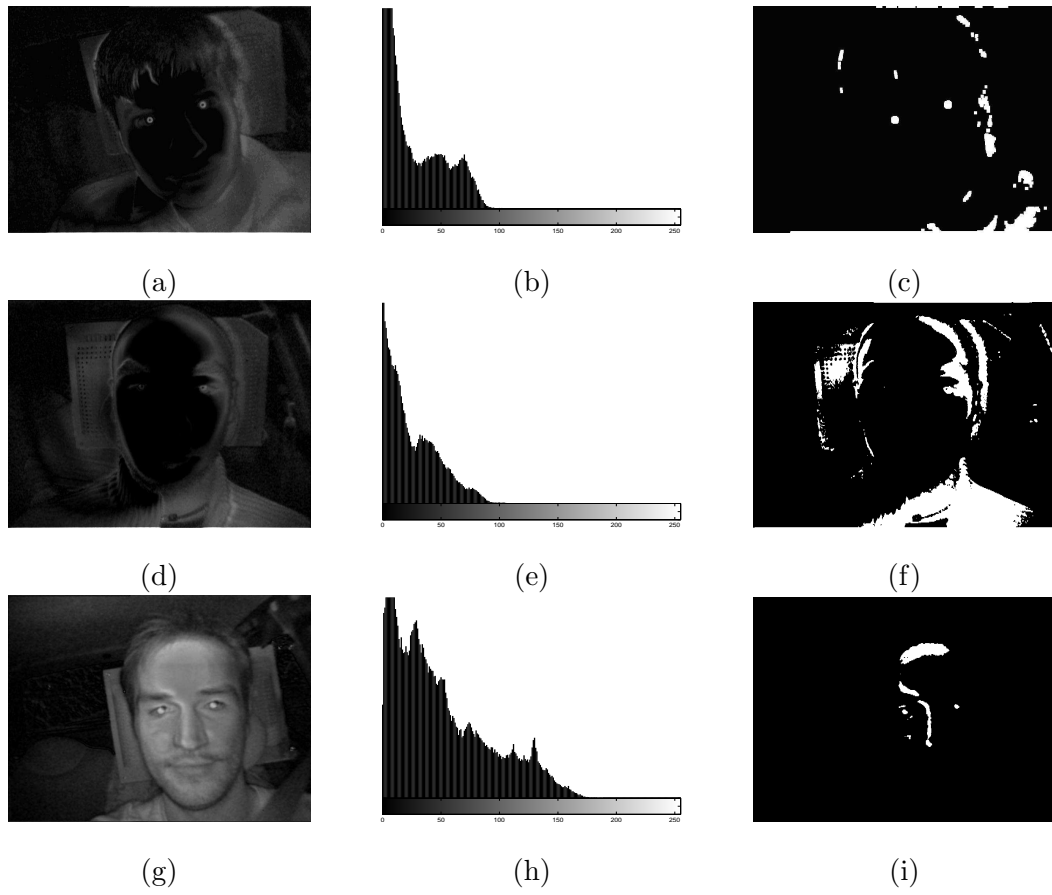


Figura 5.6: Diferencia de imágenes y su umbralización, (a-d-g) imagen diferencia, (b-e-h) histograma de la imagen diferencia, (c-f-i) imagen binarizada.

4. Calcular el promedio de los píxeles de los grupos $G1$ y $G2$ para obtener μ_1 y μ_2 , respectivamente
5. Calcular el nuevo valor de T como: $T = (\mu_1 + \mu_2)/2$
6. Repetir los pasos 3, 4 y 5 hasta que la diferencia con T en sucesivas iteraciones sea menor que un error predeterminado
7. Obtener el nuevo umbral T

En este método no se ha supuesto ninguna forma funcional del histograma, más bien, se ha tomando en cuenta el interés de resaltar los píxeles de mayor valor y excluir a los de menor valor. Ejemplos de este procedimiento se pueden observar en la figura 5.6 (c-f-i).

5.2.3. Localización de los candidatos a ojos basada en la imagen de bordes

El efecto del brillo en los ojos producto de la iluminación infrarroja es sensible a los cambios generados por las condiciones ambientales, estos cambios no son controlables y hacen que el sistema sea altamente inestable. Sin embargo, la imagen de bordes generada a partir de la imagen diferencia presenta mejores condiciones frente a los cambios externos, llegando a comportarse de forma estable en la mayoría de situaciones, y en especial cuando los ojos se encuentran semicerrados.

La imagen de bordes se obtiene utilizando el operador Sobel, seguido de un filtraje con el filtro de la mediana (figura 5.7), para nuevamente utilizar el operador de Sobel seguido de una dilatación. Recordar que el operador Sobel calcula el gradiente de la intensidad de la imagen en cada punto, resaltando los cambios de oscuro a claro; en este sentido, esta parte del algoritmo resalta las partes más claras de la imagen, en este caso, las que corresponden a los ojos. Los resultados de este procedimiento se observan en las imágenes (a-d-g) de la figura 5.8.

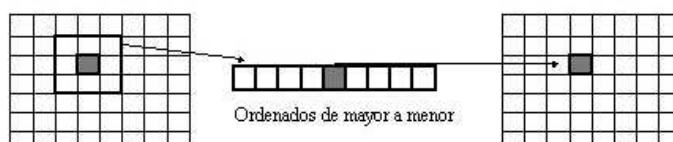


Figura 5.7: Representación del filtro de la mediana.

5.2.4. Método propuesto para umbralizar la imagen de bordes

Esta estabilidad se presenta sobre el histograma, el mismo que tiene la forma de una campana que concentra la mayoría de sus valores en los tonos oscuros de la escala de grises, dejando una mínima cantidad de píxeles para los tonos blancos (ver figura 5.8 (b-e-h)).

La forma del histograma puede ser modelada por la función de densidad de probabilidad $\text{Gamma}(\alpha, \beta)$; distintos valores de los parámetros α y β generan distintas formas de

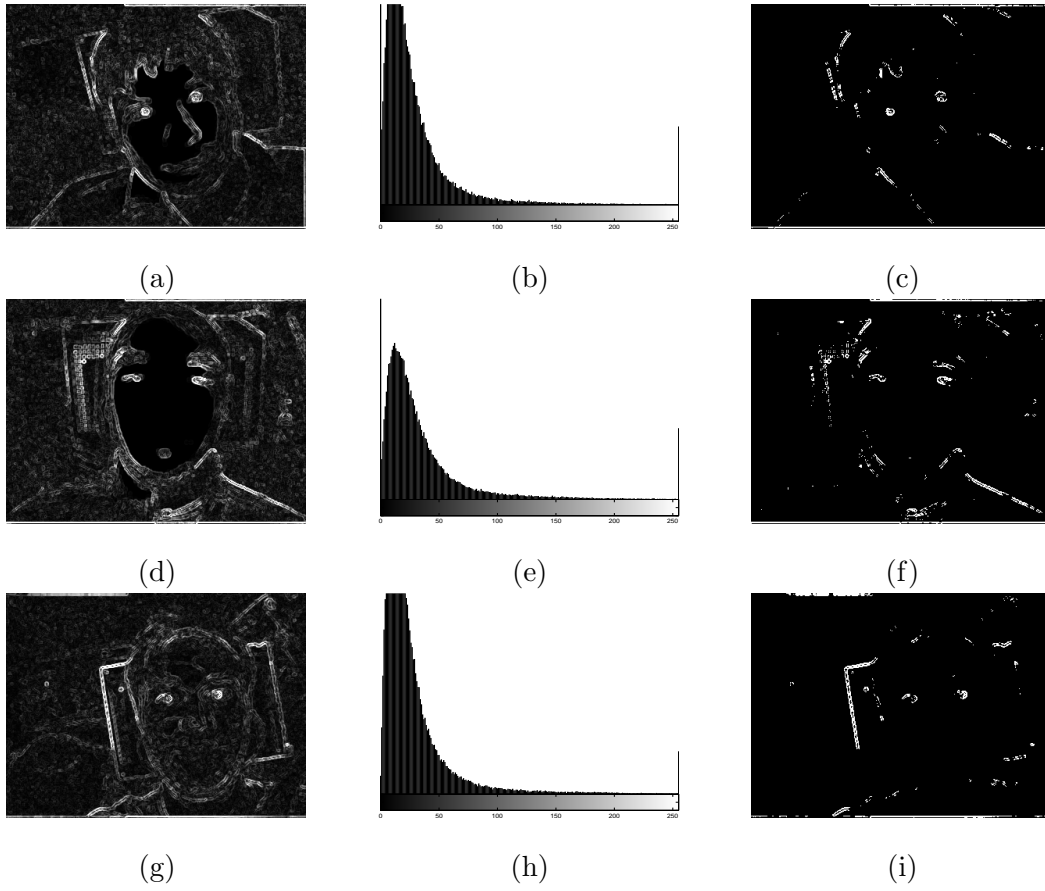


Figura 5.8: Imagen de bordes de la imagen diferencia (a-d-g), su histograma (b-e-h) y su imagen binarizada (c-f-i).

esta función de distribución. Su forma funcional es la siguiente [128].

$$f(x; \alpha, \beta) = \begin{cases} x^{\alpha-1} \frac{e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)} & \text{si } x \geq 0 \\ 0 & \text{caso contrario} \end{cases} \quad (5.5)$$

donde $\Gamma(\alpha) = (\alpha - 1)!$ y el símbolo $!$ representa el factorial de un número. Para estimar los parámetros que generan la forma del histograma se utiliza el método de momentos, el mismo que viene dado por las siguientes ecuaciones:

$$\hat{\alpha} = \frac{E(X)^2}{Var(X)} \quad (5.6)$$

$$\hat{\beta} = \frac{Var(X)}{E(X)} \quad (5.7)$$

Aplicando las ecuaciones (5.6) y (5.7) sobre las imágenes de la figura 5.8 se obtienen los siguientes resultados, así, para la imagen (a) se generan los siguientes parámetros $\hat{\alpha} = 0,5926$, $\hat{\beta} = 48,4139$, para la imagen (d) se tienen $\hat{\alpha} = 0,5736$, $\hat{\beta} = 54,4592$, y para la imagen (g) se producen los siguientes resultados $\hat{\alpha} = 0,7115$, $\hat{\beta} = 40,4795$. Al 90 % se obtiene los siguientes valores 200, 205, 170, respectivamente.

El procedimiento propuesto para umbralizar la imagen de bordes consiste en tomar el intervalo que contiene el 90 % de los datos, y así obtener el valor del umbral. Con éste umbral se consigue otro conjunto de cuerpos de entre los cuales se encuentran los ojos, en la mayoría de casos, estos cuerpos corresponden plenamente con los obtenidos en la sección anterior, pero cuando los ojos se encuentra cerrados o semicerrados, los resultados son prometedores. En las imágenes de la figura 5.8 (c-f-i) se aprecian varios ejemplos de lo expuesto.

5.2.5. Verificación y localización de los ojos

La estrategia planteada para localizar los ojos se basa en la información proveniente de la imagen binaria (generada a partir de la imagen diferencia y su imagen de bordes, umbralizadas), en conjunto con ciertas operaciones para ir eliminando sistemáticamente los cuerpos que no corresponden a los ojos.

En la imagen binaria, donde se localizan varios cuerpos de diferentes tamaños y formas se presentan los escenarios que se ilustraran a través de los ejemplos:

- El primer escenario corresponde al caso donde la imagen binarizada contiene únicamente dos cuerpos, que resultan ser los ojos. En este caso los ojos se encuentran directamente sin ningún otro procedimiento posterior.
- El segundo escenario corresponde al caso donde la imagen binarizada contiene un solo cuerpo, en esta situación es necesario descartar la imagen y continuar con la siguiente o tratar de detectar el rostro directamente.
- El tercer escenario corresponde al caso más realista, donde la imagen binarizada contiene varios cuerpos de entre los cuales se encuentran los ojos, en este caso, es necesario eliminar los cuerpos que no corresponde a los ojos hasta lograr aislarlos.

5.2 Detección de los ojos

Es así que se llega a la conclusión que es necesario desarrollar una estrategia que permita eliminar los cuerpos que no corresponden a los ojos. Para ello, se empleará un método que permita clasificar entre un ojo y un no-ojo. En este sentido se presenta la siguiente estrategia que consta de dos pasos:

1. Eliminación de cuerpos basada en restricciones geométricas
2. Clasificación basada en el clasificador SVM

5.2.6. Eliminación basada en restricciones geométricas

La primera estrategia para eliminar los cuerpos que no corresponden a los ojos depende de las propiedades geométricas de dichos cuerpos.

La forma y tamaño de los cuerpos que se forman en la imagen binarizada dependen principalmente de dos factores: (i) de la distancia entre el conductor y la cámara y, (ii) de la intensidad de la iluminación. Generalmente, los cuerpos que corresponden a las pupilas son circulares, en función de ésta propiedad se imponen las restricciones geométricas para eliminar una buena cantidad de falsos candidatos.

- En primer lugar se eliminan los cuerpos que superan un determinado umbral de área, éste umbral ha sido elegido experimentalmente.
- Luego, se eliminan los cuerpos que superan un determinado umbral del cociente generado por el perímetro del cuerpo y su área, cuyo valor también ha sido elegido experimentalmente.
- En tercer lugar, se eliminan los cuerpos que presentan una relación ancho-alto desproporcionada, dando prioridad para ser eliminados los cuerpos que presentan mayor alto que ancho.

Con esto se consigue eliminar una buena cantidad de falsos candidatos a ojos, pero aún sobran muchos cuerpos que no corresponden a los ojos pero que cumplen las restricciones anteriormente impuestas, por ello, se procede con el método que viene a continuación.

5.2.7. Eliminación basada en el clasificador SVM

Con el proceso anterior se consigue eliminar una gran cantidad de cuerpos que no corresponden a los ojos, sin embargo, aún no se han localizado plenamente. A pesar de estos inconvenientes, este método es bastante efectivo para reducir el tiempo de cómputo ya que permite eliminar cuerpos extraños rápidamente antes de pasar al siguiente paso, que consiste en utilizar un método de clasificación.

Para eliminar los cuerpos que no corresponden a los ojos se ha utilizado un método de clasificación no supervisado. Partiendo de esta necesidad se ha vuelto a utilizar el clasificador *SVM* para clasificar los cuerpos como ojos y no-ojos.

Antes de proseguir con el desarrollo del método se introduce la definición de región de interés (RI) del ojo como función de área del cuerpo generado en la imagen binarizada.

5.2.7.1. Definición : Región de interés del ojo

Para enmarcar al ojo suponiendo que se conoce el área de su pupila, se introduce la región de interés como función del área del cuerpo generado en la imagen binarizada que se expresa a través de la siguiente ecuación.

$$A_{ojo} = \lambda A_{pupila} \quad (5.8)$$

λ es una constante que toma distintos valores de acuerdo al tamaño del área de la pupila, en esta tesis, este valor se categoriza en cinco casos elegidos experimentalmente. En la figura 5.9 se observa un ejemplo.

Para lo que sigue en las siguientes secciones, se considerará el área que genera la ecuación (5.8) como el cuerpo objeto de estudio.

5.2.7.2. Extracción de características basada en el filtro de Gabor

Para entrenar el clasificador SVM se construyó una base de datos de entrenamiento con imágenes de ojos abiertos e imágenes de no-ojos, en ambos casos, bajo condiciones de iluminación infrarroja. Esta base consta de 366 ojos abiertos y de 595 no-ojos. El tamaño de cada una de las imágenes es 40 píxeles de ancho por 40 píxeles de alto.

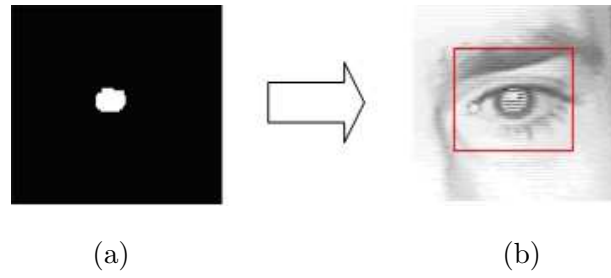


Figura 5.9: Estimación de la región de interés del ojo a partir del área de la pupila, (a) área de la pupila en la imagen binarizada, (b) región de interés en la imagen con iluminación infrarroja.

Para entrenar SVM se utilizó el filtro de Gabor [54] con una escala $\{0\}$ y cuatro orientaciones $\{0, \pi/4, \pi/2, 3\pi/4\}$, estos parámetros fueron obtenidos experimentalmente. En la figura 5.10 se muestran dos ejemplos de este proceso, el primero corresponde al caso de un ojo y el segundo corresponde al de un no-ojo.

En los ejemplos utilizados para entrenar SVM se observa que al aplicar el filtro de Gabor se consigue resaltar la forma del ojo, controlando en cierto sentido la iluminación que corrompe la forma intrínseca del mismo. Esto mejora sustancialmente la calidad de las imágenes de entrenamiento.

El vector de características se encuentra formado por el nivel de grises de las cuatro imágenes entregadas por el filtro de Gabor, motivo por el cual se tiene un vector de dimensión 6.400, que corresponde al producto del ancho por el alto de la imagen de entrada.

5.2.7.3. Reducción de la dimensión del vector de características: EigenOjosGabor

Para reducir el tamaño del vector de características del ojo (imagen del ojo generado con el filtro de Gabor) sin perder información relevante que lo represente se ha decidido utilizar el conocido método PCA (Principal Component Analysis). Con este método se construye el denominado espacio de ojos llamado EigenOjosGabor [99], [6], más concretamente, se puede decir que PCA es un método de proyección que mapea un conjunto de datos en un espacio de menor dimensión, a través de una transformación lineal que

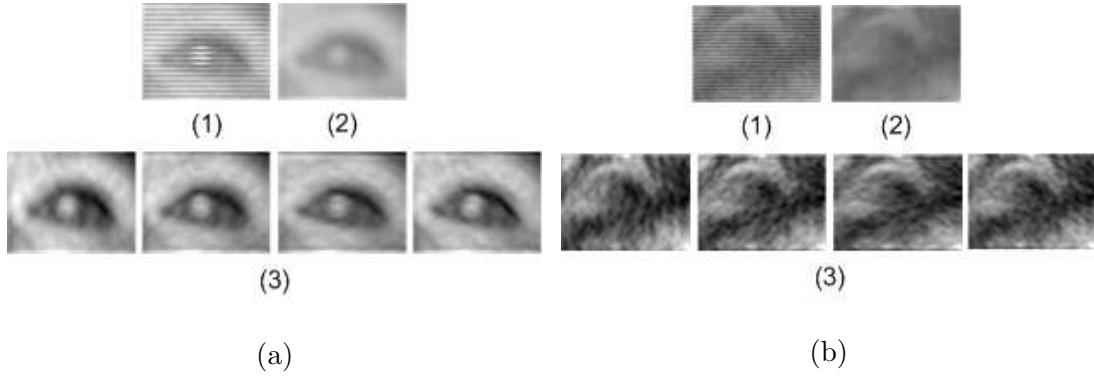


Figura 5.10: Filtro de Gabor aplicado a las imágenes de entrenamiento de SVM con una escala $\{0\}$ y cuatro orientaciones $\{0, \pi/4, \pi/2, 3\pi/4\}$: (a) imagen de un ojo, (b) imagen de un no ojo. En los dos casos (1) corresponde a la imagen original de entrada, (2) a su imagen filtrada con el filtro de la media, y (3) corresponde a las respuestas del filtro de Gabor.

mantiene la mayor cantidad de información (en el sentido de retener la mayor varianza de los datos proyectados) de entre todas las proyecciones lineales [144] [84].

5.2.7.4. Cálculo de los EigenOjosGabor

Se dispone de un conjunto de imágenes de ojos $\{I_1, \dots, I_M\}$. La imagen del promedio del ojo es definida por $\Psi = \frac{1}{M} \sum_{j=1}^M I_j$. Cada imagen difiere del promedio por el vector $\Phi_i = I_i - \Psi$. El objetivo es buscar un conjunto de vectores linealmente independientes u_j , $j = 1, \dots, M$, que mejor describa la distribución de los datos. El k -ésimo vector se elige tal que el valor de la ecuación (5.9) sea máximo dada la restricción (5.10).

$$\lambda_k = \frac{1}{M} \sum_{j=1}^M u_k^T \Phi_j \quad (5.9)$$

$$u_l^T u_k = \delta_{lk} = \begin{cases} 1 & \text{si } l = k \\ 0 & \text{si } l \neq k \end{cases} \quad (5.10)$$

Los vectores u_k y los escalares λ_k son los eigenvectores y eigenvalores, respectivamente, de la matriz de covarianza:

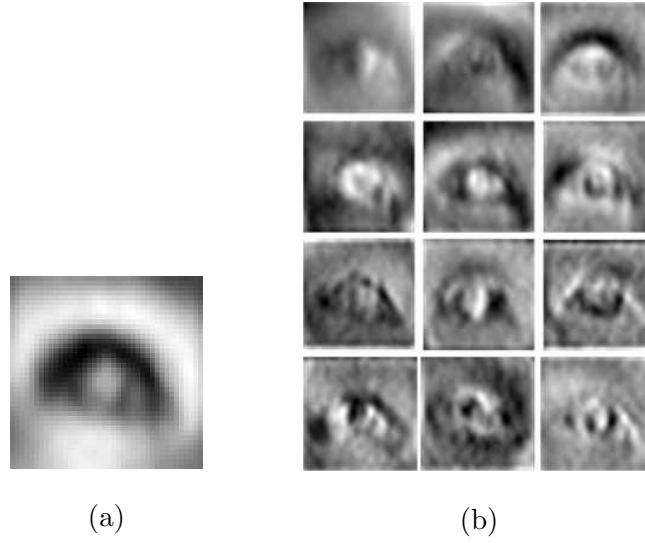


Figura 5.11: Resultados de PCA, (a) Imagen promedio de los ojos, (b) ejemplos de los primeros 12 eigenojosgabor.

$$C = \frac{1}{M} \sum_{j=1}^M \Phi_j \Phi_j^T = A A^T \quad (5.11)$$

donde la matriz $A = [\Phi_1, \dots, \Phi_M]$.

En la figura 5.11(a) se observa la imagen promedio, y en la figura 5.11 (b) se presentan los primeros 12 eigenojosgabor obtenidos bajo este análisis.

De acuerdo a la figura 5.12 (b) se toman 136 eigenvectores para formar el espacio EigenOjosGabor, esto significa que se ha logrado una reducción del 72 % en las dimensiones del vector de características manteniendo el 95 % de la información. Lo anterior indica que al proyectar un nuevo ojo sobre el espacio de EigenOjosGabor se obtiene un vector que conserva el 95 % de la información.

Para finalizar el estudio de los EigenOjosGabor, en la figura 5.13 (a) se presenta el diagrama de caja (boxplot) de las distancias de los ojos y no-ojos proyectados sobre este espacio, se puede observar claramente que la media de la distancia de los ojos es menor que la media de la distancia de los no-ojos, lo que en principio anticipa un proceso de clasificación potente.

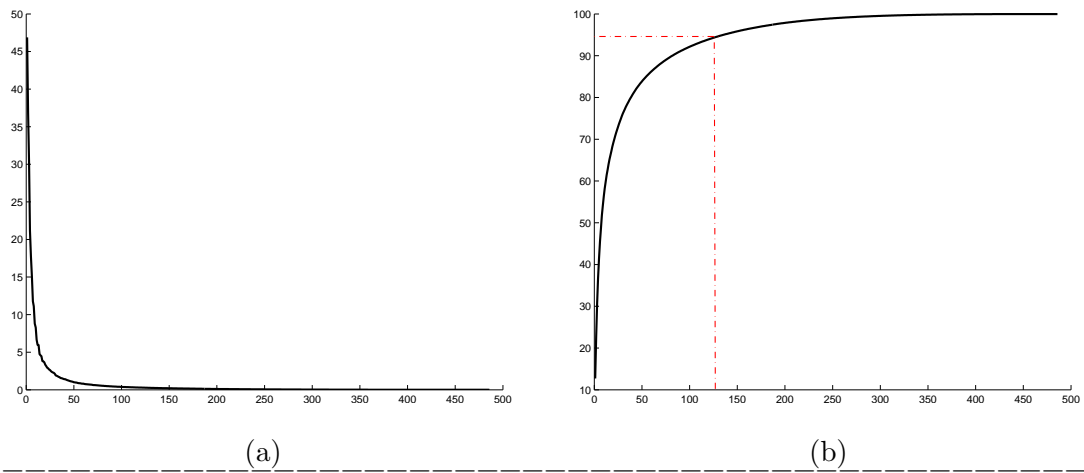


Figura 5.12: Representación de los eigenvalores del PCA de los ojos (a) eigenvalores, (b) eigenvalores acumulados.

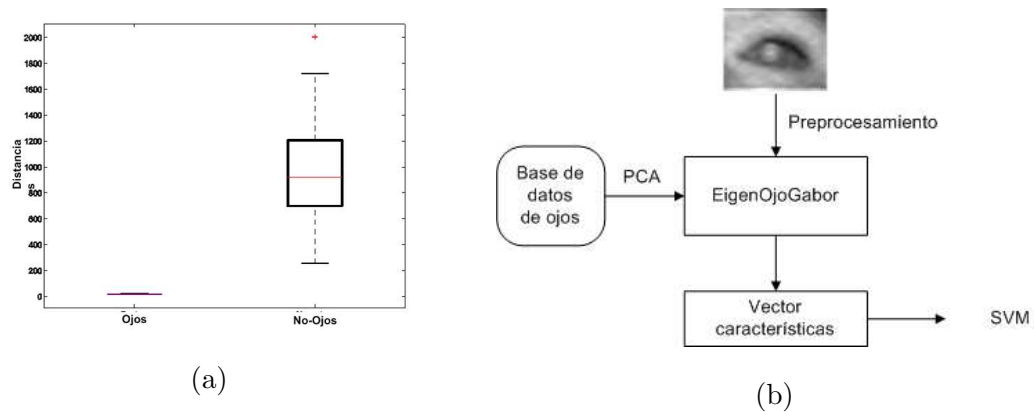


Figura 5.13: Proyecciones en el espacio Eigenrosto, (a) comparación de las distancias entre rostros y no rostros proyectados sobre el Eigenrosto, (b) extracción de características por medio de PCA.

5.2.7.5. Entrenamiento de SVM

Una vez obtenida la base de datos de ojos y no-ojos debidamente procesada, y siguiendo el esquema de la figura 5.13 (b), los parámetros de entrenamiento de SVM son los siguientes:

- C : 100

5.3 Detección del rostro

- γ : 0,128
- *Kernel*: RBF

Con ellos se alcanza un tasa de clasificación del 93 % sobre una base de datos de prueba que consta de 220 ojos y 500 no-ojos.

5.2.8. Resultados experimentales en la localización de los ojos

En la figura 5.14 se aprecian los resultados generados con el procedimiento de localización de los ojos. Este método trabaja bastante bien pero tiene un inconveniente, solo es capaz de clasificar los ojos cuando ellos se encuentran abiertos.



Figura 5.14: Ejemplos de la detección de los ojos bajo el método propuesto.

Por lo anterior, este método es adecuado para localizar ojos abiertos bajo iluminación infrarroja e inicializar el módulo de seguimiento de los ojos.

Recordar que en este caso SVM ha sido utilizado de forma diferente al caso del espectro visible, aquí ha servido para clasificar entre un ojo y un no-ojo, bajo la influencia de la iluminación infrarroja cercana.

5.3. Detección del rostro

En este caso, para detectar el rostro se ha decidido proponer una nueva estrategia, distinta a la utilizada en el espectro visible, cuya base fue el algoritmo de Viola & Jones

[120], [118], [119]. Recordando que detectarlo sobre una imagen en el espectro visible es dificultoso [43], el trabajo se complica mucho más en el caso de imágenes capturadas bajo la iluminación infrarroja.

Por lo tanto, el método para la detección del rostro, que se presenta en esta sección, trabaja directamente sobre las imágenes capturadas bajo la influencia de la iluminación infrarroja, ver figura 5.15, y consta de las siguientes etapas:

- Modificación del contraste
- Cálculo de la función de varianza vertical
- Estimación del rostro a partir de la posición de los ojos

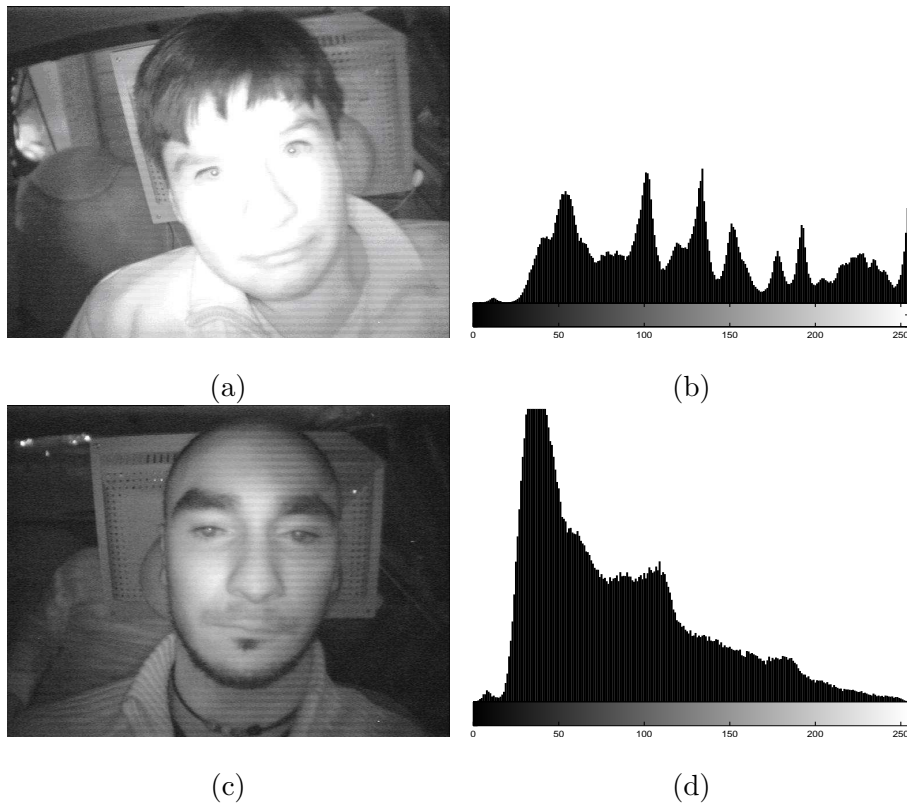


Figura 5.15: Imágenes generadas con el sistema de iluminación infrarroja instalado sobre el vehículo IvvI: (a) y (c) imágenes de conductores, (b) y (d) sus histogramas.

5.3.1. Modificación del contraste

El sistema de iluminación diseñado para este trabajo genera un brillo que se manifiesta en especial sobre el cuerpo (rostro) del conductor. Este exceso de iluminación se observa al revisar la forma del histograma en las imágenes de la figura 5.15 (b) y (d).

Experimentalmente se ha comprobado que al modificar el contraste a través de la ecuación (5.12) se favorecen los píxeles que tienen mayor valor en perjuicio de los que tienen menor valor dentro de la imagen [28],[66]; lo que al final permite mejorar el método de localización de la posición del cuerpo.

$$J(x, y) = \frac{I(x, y)^3}{255^2 N} \quad (5.12)$$

donde $I(x, y)$ representa el valor del pixel en la escala de grises en la posición (x, y) , N es la dimensión de la imagen, es el producto de su ancho por su alto, y la fracción $\frac{1}{255^2}$ es el factor de normalización en el intervalo $[0, 255]$.

Al modificar el contraste se obtienen imágenes similares a las presentadas en la figura 5.16, donde se puede ver que el cuerpo del conductor ha quedado separado de su fondo, a la vez que sus histogramas presentan un mejor comportamiento, que incluso se puede decir que tienen una distribución uniforme.

5.3.2. Funciones de proyección y varianza vertical

A partir de la imágenes (a) y (c) de la figura 5.16 se calcula la proyección vertical (también denominada proyección integral vertical), usando la ecuación (5.13) [55], [63], [100].

$$Pr(x) = \frac{1}{h} \sum_{y=1}^h I(x, y) \quad (5.13)$$

Los resultados se pueden mejorar al utilizar la ecuación (5.14), que es la función de varianza de la proyección vertical.

$$\sigma_V^2(x) = \frac{1}{h} \sum_{y=1}^h [I(x, y) - Pr(x)]^2 \quad (5.14)$$

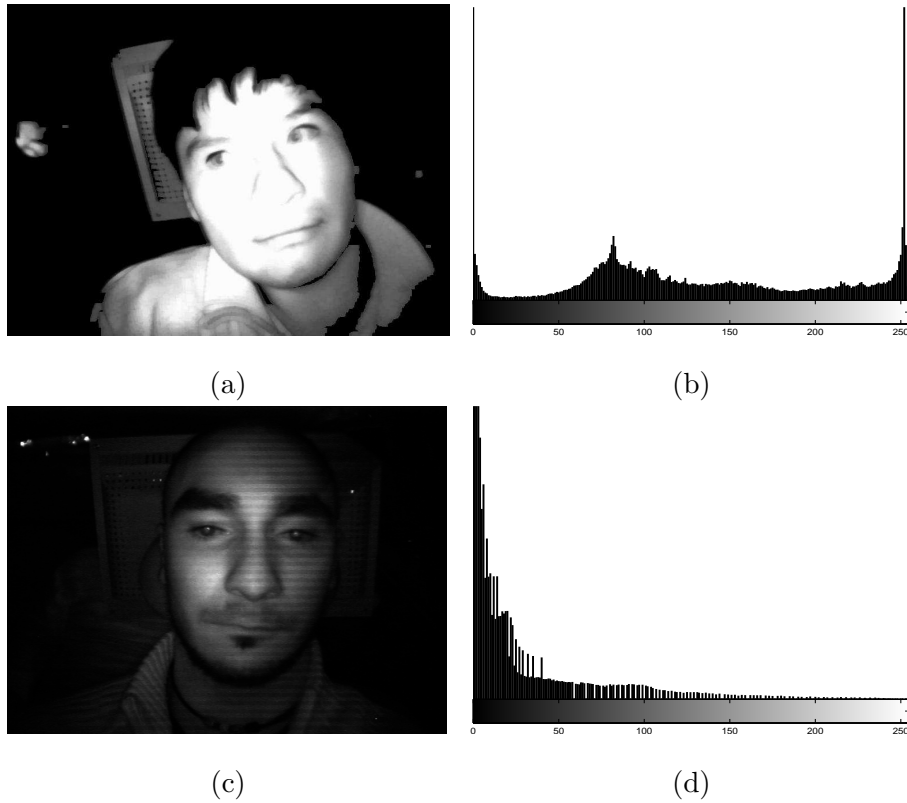


Figura 5.16: Modificación del contraste: (a) y (b) imágenes con el contraste modificado, (c) y (d) sus respectivos histogramas.

En las dos ecuaciones h representa el alto de la imagen, $I(x, y)$ es el valor del píxel en la posición (x, y) . La diferencia entre las dos ecuaciones es que la segunda genera menor dispersión en la información debido a que toma en cuenta el valor promedio de los datos. Los gráficos de estas ecuaciones se muestran en la figura 5.17 (a), (b), (d) y (e).

El valor máximo de la función de proyección permite construir una región que contiene al cuerpo del conductor, donde cuerpo se entiende al objeto formado por la cabeza y el tronco. El valor máximo de la proyección es comprobado con la primera derivada que equivale al mínimo valor que ésta adquiere. Se observa en las imágenes de la figura 5.17 (c) y (f), que la proyección vertical es prácticamente simétrica respecto al punto más alto (punto **B**); esto contribuye significativamente a estimar la zona donde se localiza el rostro (cabeza) del conductor, entre los puntos **A** y **C**.

Para finalizar esta sección, en la figura 5.18 se presentan dos ejemplos de la detección

5.3 Detección del rostro

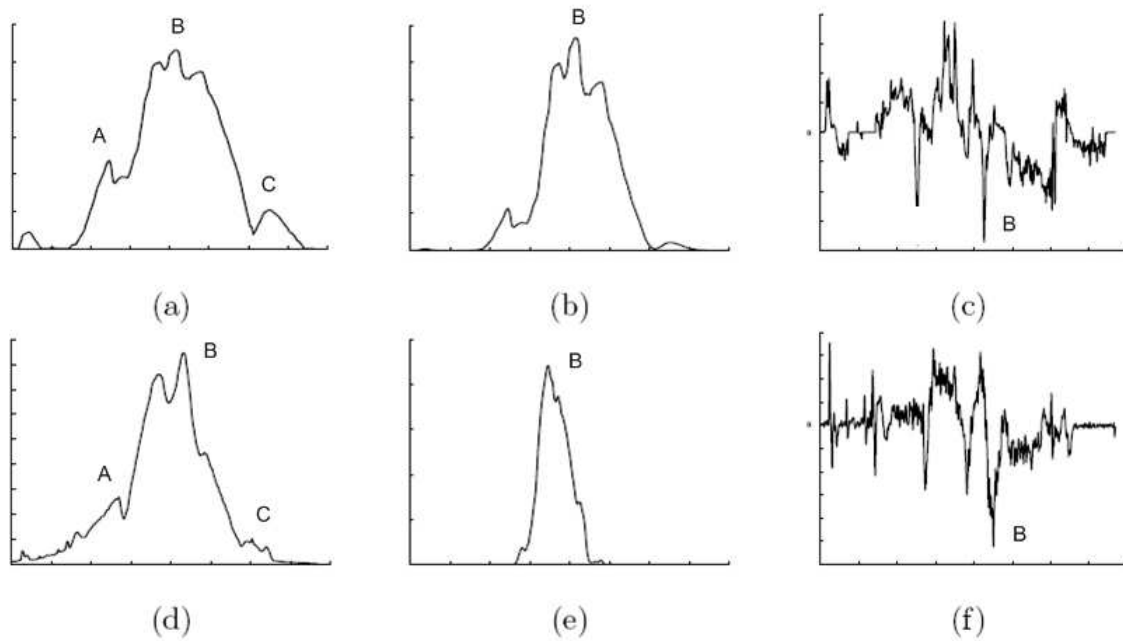


Figura 5.17: Proyección vertical para detectar el cuerpo y el rostro del conductor basados en las imágenes de la figura 5.16, (a) y (d) proyección vertical, (b) y (e) varianza de la proyección vertical y (c) y (f) derivadas de la función de proyección vertical



Figura 5.18: Resultados obtenidos en la detección del rostro y los ojos.

del rostro y los ojos sobre dos imágenes tomadas en el vehículo IvvI. En ellos se localiza el rostro utilizando la información de la figura 5.17, en verde la proyección vertical y en rojo la varianza vertical.

5.3.3. Estimación del rostro a partir de la posición de los ojos

De acuerdo al diagrama de la figura 5.1, la información del rostro es utilizada para restringir la zona de búsqueda de los ojos, y viceversa, con la información de la posición de los ojos se estima la zona del rostro.

Una vez que se ha localizado la posición exacta de los dos ojos se procede a estimar los parámetros del rostro, esto es, la posición y los ejes de la elipse que lo contienen [86], [115]. La figura 5.19 muestra los parámetros que sirven para el modelo que se ha propuesto.

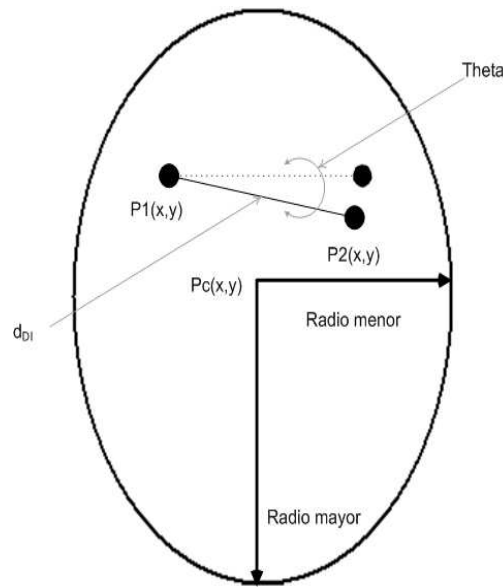


Figura 5.19: Modelo paramétrico del rostro a partir de la posición de los ojos.

$$d_{DI} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (5.15)$$

$$\theta = \text{atan} \left(\frac{y_2 - y_1}{x_2 - x_1} \right) \quad (5.16)$$

Los radios mayor y menor y el centro son función de la distancia que separa a los ojos y vienen dados por las ecuaciones (5.17), (5.18) y (5.19), respectivamente.

$$R_1 = 1,5d_{DI} \quad (5.17)$$



Figura 5.20: Resultados de la estimación del rostro a partir de la posición de los ojos en imágenes tomadas sobre IvvI.

$$R_2 = d_{DI} \quad (5.18)$$

$$(x_c, y_c) = (x_1 + 0,5d_{DI}, y_1 + 0,3d_{DI}) \quad (5.19)$$

Con estas ecuaciones se logra enmarcar plenamente la zona que corresponde al rostro. En la figura 5.20 se aprecian dos ejemplos donde se observa la elipse que encuadra el rostro de dos conductores.

Para evitar falsas detecciones del rostro, se restringe la zona de búsqueda de los ojos a la zona generada con la función de proyección de varianza vertical.

5.4. Enfoque bayesiano para el seguimiento de los ojos y el rostro

La necesidad de mantener localizados a los ojos y al rostro en cada instante hace que sea necesario incorporar un módulo de seguimiento. En este caso se empieza por describir el seguimiento de los ojos, luego se continua con el seguimiento del rostro, en ambos casos, basados nuevamente en el algoritmo Condensation.

5.4.1. Seguimiento de los ojos

En este caso, el estado (posición y velocidad) y la observación (posición) de cada ojo vienen dados por las ecuaciones (5.20) y (5.21), respectivamente.

$$x_t = (x, y, u, v) \quad (5.20)$$

$$z_t = (x, y) \quad (5.21)$$

donde (x, y) representa la posición del centro del ojo y (u, v) es su velocidad en las direcciones x e y , respectivamente.

El modelo de transición, que relaciona el estado presente con el estado inmediatamente anterior viene dado por la ecuación (5.22), que representa un proceso autoregresivo de orden uno $AR(1)$.

$$x_t = Ax_{t-1} + \xi_t \quad (5.22)$$

y el modelo de la medida

$$z_t = Hx_t + \eta_t \quad (5.23)$$

nuevamente, ξ_t y η_t son ruidos blancos que representan las perturbaciones del sistema, A es la matriz que relaciona los estados en dos instantes de tiempo y H es la matriz que relaciona el estado con la medida [5].

5.4.1.1. Evaluación de la función de densidad

Para evaluar los pesos $\pi_t^{(n)}$ se sigue una doble estrategia que consiste en utilizar la distancia euclídea generada por la medida de los cuerpos que se localizan en la zona de influencia generada en el instante anterior por la probabilidad de cada pixel generado en la imagen diferencia.

En el primer caso se tiene:

$$f_1(z) = \frac{1}{k_1} \exp(-k_1 * \|s - z\|^2) \quad (5.24)$$

5.4 Enfoque bayesiano para el seguimiento de los ojos y el rostro

k_1 representa la dispersión de las partículas respecto a la medida, la medida $z = z(x, y)$ se extrae de los cuerpos generados en las imágenes binarizadas, en este caso, corresponde al centro de gravedad de dichos cuerpos, y $s = s(x, y)$ es la posición en el punto (x, y) extraída de la partícula $s_t^{(n)}$.

Para evaluar la última probabilidad se construye la función de densidad triangular.

$$f_2(s) = \begin{cases} \frac{m}{k_2} s & \text{si } s \in [0, 255] \\ 0 & \text{caso contrario} \end{cases} \quad (5.25)$$

el valor $m = \frac{1}{255}$ es la pendiente de la función triangular generada para este proceso, k_2 es un factor de normalización que garantiza que se trata de una función de densidad de probabilidad, y $s = s(I_D(x, y))$ es el valor del pixel en la posición extraída de la partícula $s_t^{(n)}$ sobre la imagen diferencia I_D en el instante t . En la figura 5.21 se observa la forma de la función de densidad triangular, donde los píxeles de mayor valor aportan con mayor información, es decir, presentan la probabilidad más alta de pertenecer al ojo bajo el efecto de la iluminación infrarroja.

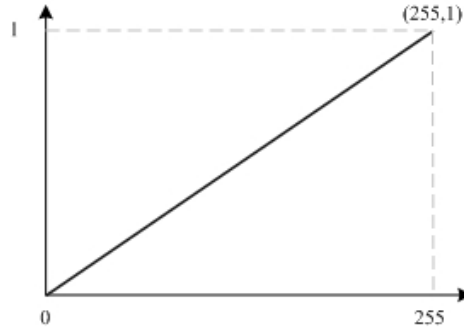


Figura 5.21: Representación de la función de densidad triangular para evaluar los pesos del algoritmo Condensation.

En resumen se tiene que la evaluación de los pesos de la función de distribución a posteriori viene dada por la ecuación (5.26).

$$\pi_t^{(n)} = f_1(z) * f_2(s) \quad (5.26)$$

5.4.1.2. Resultados experimentales del seguimiento de los ojos

Para iniciar la ejecución de CA se debe conocer la distribución inicial del estado, es decir, conocer $p(x_0)$; en este caso, se utiliza la distribución normal de parámetros X_0, Σ_0 , es decir, $p(x_0) \equiv N(X_0, \Sigma_0)$, donde el estado inicial puede ser descrito por (5.27) cuando se conoce el estado en dos frames consecutivos t y $t+1$, y la matriz de varianza se establece por (5.28).

$$\begin{aligned} x_0 &= x_{t+1} \\ y_0 &= y_{t+1} \\ u_0 &= x_{t+1} - x_t \\ v_0 &= y_{t+1} - y_t \end{aligned} \tag{5.27}$$

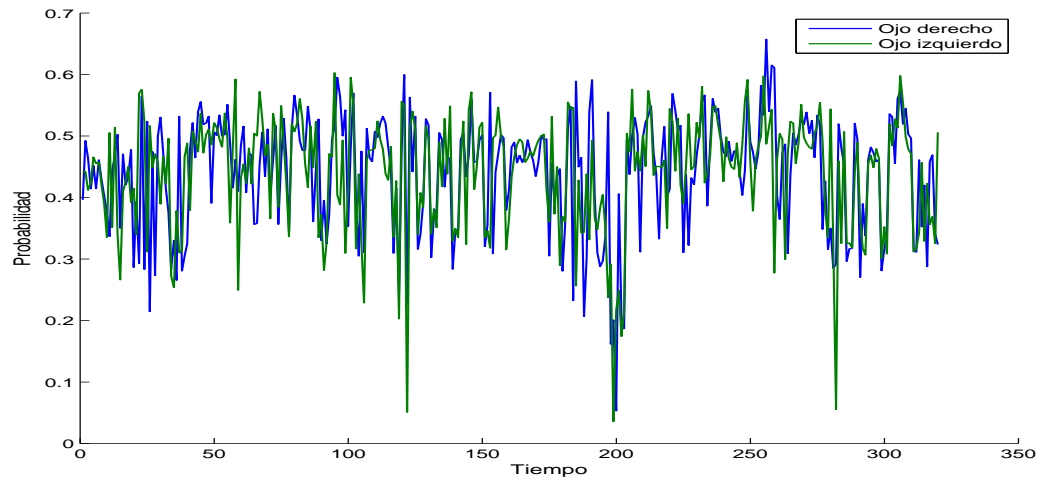
$$\Sigma_0 = \begin{bmatrix} 100 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 \\ 0 & 0 & 25 & 0 \\ 0 & 0 & 0 & 25 \end{bmatrix} \tag{5.28}$$

Con todo esto, los resultados que se obtienen se resumen en los valores de la tabla 5.1.

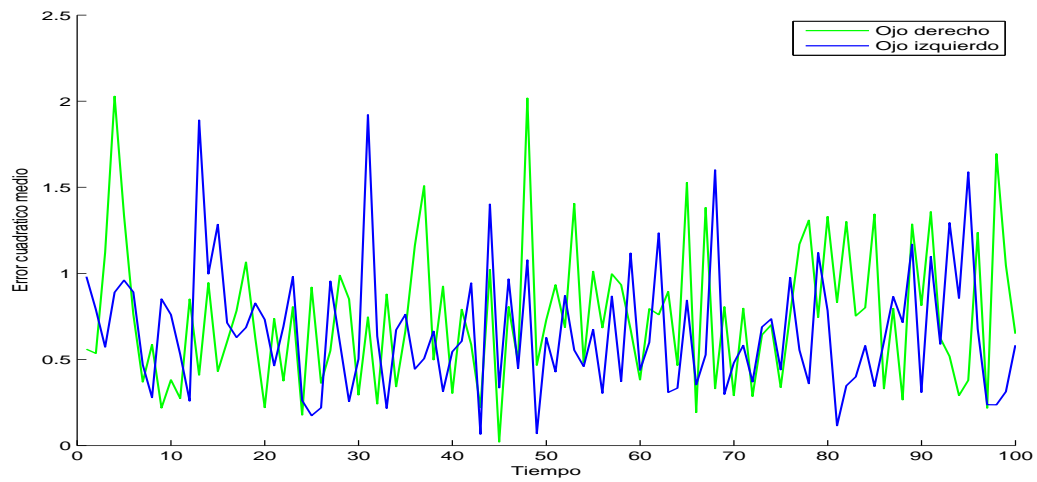
	Número de imágenes	Número de fallos	Aciertos (%)
Secuencia 1	800	14	98,25
Secuencia 2	646	29	95,51
Secuencia 3	600	24	96,00
Secuencia 4	324	9	97,22

Tabla 5.1: Resultados del seguimiento de los ojos bajo iluminación infrarroja generado por el algoritmo Condensation sobre varias secuencias de imágenes de conductores.

En la figura 5.22 (a) se muestran los valores calculados de la función de probabilidad de las observaciones, y en la figura 5.22 (b) se presentan los errores cometidos en los ojos derecho e izquierdo, respectivamente, sobre una secuencia formada por 324 fotogramas.



(a)



(b)

Figura 5.22: Resultados del algoritmo Condensation en el seguimiento los ojo, (a) probabilidad de las observaciones y (b) errores generados en el seguimiento de los ojos derecho e izquierdo, respectivamente.

5.4.2. Seguimiento del rostro

La fuerte dependencia de la iluminación infrarroja para detectar la posición de los ojos hace que el sistema sea inestable tanto en la detección como en el seguimiento, especialmente, en el caso de que los ojos permanezcan cerrados durante mucho tiempo (adormecimiento), ver la figura 5.23. Para corregir esta clase de inconvenientes se incorpora el módulo de seguimiento del rostro.

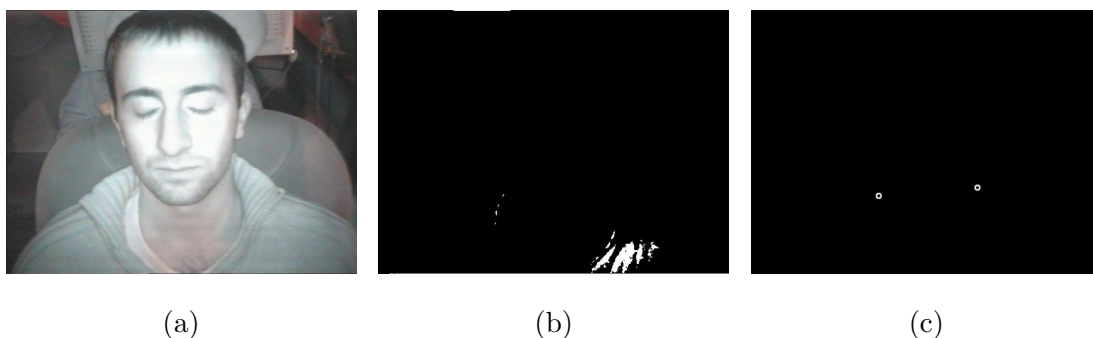


Figura 5.23: Ejemplo donde los ojos se encuentran cerrados, (a) imagen bajo iluminación infrarroja, (b) imagen binaria obtenida a partir de la imagen diferencia, (c) imagen binaria obtenida a partir de la imagen de bordes de la imagen diferencia.

El módulo de seguimiento que se presenta en este caso es similar al del espectro visible, por ello, se describe el entrenamiento de las redes neuronales utilizado para tal propósito.

5.4.2.1. Entrenamiento de la red neuronal artificial

Al igual que en el caso del espectro visible, en este caso nuevamente se ha decidido entrenar una red neuronal con imágenes tomadas bajo iluminación infrarroja.

El procedimiento es similar al caso anterior, por lo que no se entrará en mayores detalles, salvo en la información concerniente a las imágenes de rostros y no-rostros que fueron preprocesadas con el filtro de Gabor bajo una escala $\{0\}$ y dos orientaciones $\{\pi/4, \pi/2\}$. El preproceso ha permitido resaltar sustancialmente las características que mejor representan al rostro y lo diferencian de un no-rostro [83].

Para este proceso se ha construido una base de datos de entrenamiento que consta de

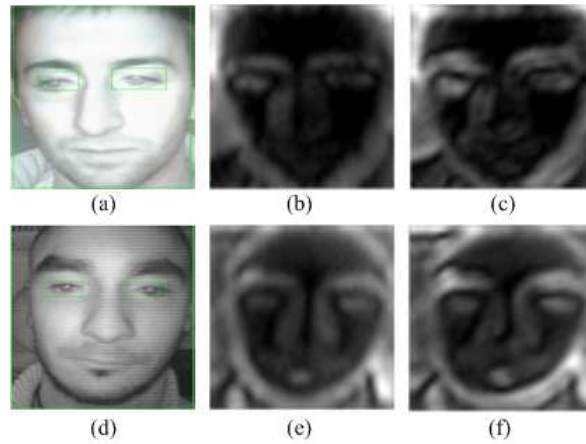


Figura 5.24: Ejemplos de rostros de la base de datos de entrenamiento (a,d), y sus respuestas al filtro de Gabor (b,c,e,f).

150 rostros y 200 no-rostros provenientes de imágenes generadas con el sistema de visión infrarrojo instalado en el vehículo Ivvi, ejemplos de esta base se observan en la figura 5.24.

Es así que los parámetros de entrenamiento de la red, luego de varios experimentos, son los siguientes:

- Número de neuronas de entrada: 1.600
- Número de neuronas de salida: 1
- Número de capas: 3
- Número de neuronas ocultas: 90
- Error deseado: 1×10^{-4}
- Función de activación: sigmoid

Con estos parámetros, la red ha alcanzado un 94 % de clasificación sobre una base de datos de prueba que consta de 1.000 rostros y 1.500 no-rostros. La curva ROC (Receiver Operator Curve) del conjunto de prueba se observa en la figura 5.25, donde se nota que la clasificación bajo el filtro de Gabor es más potente que sin él.

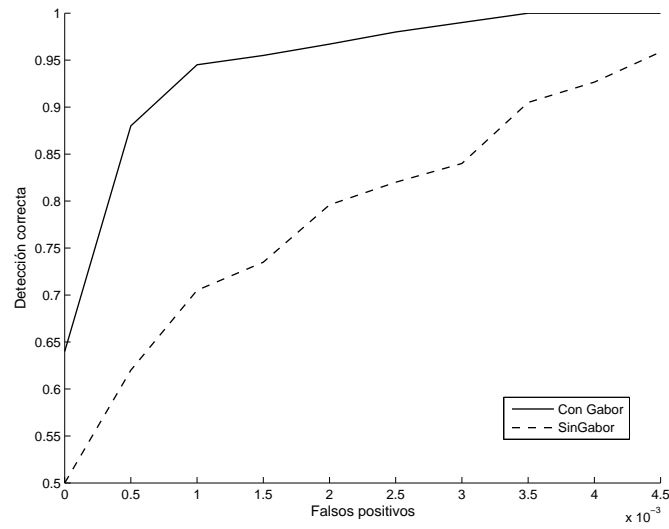


Figura 5.25: Curvas ROC generadas sobre la base de datos de prueba de rostros con y sin el filtro de Gabor.

5.4.2.2. Evaluación de la función de densidad

En este caso la evaluación de la función de densidad de las observaciones viene dada por la misma fórmula que en el caso del espectro visible. En la figura 5.26 se presenta las probabilidades obtenidas en este caso.

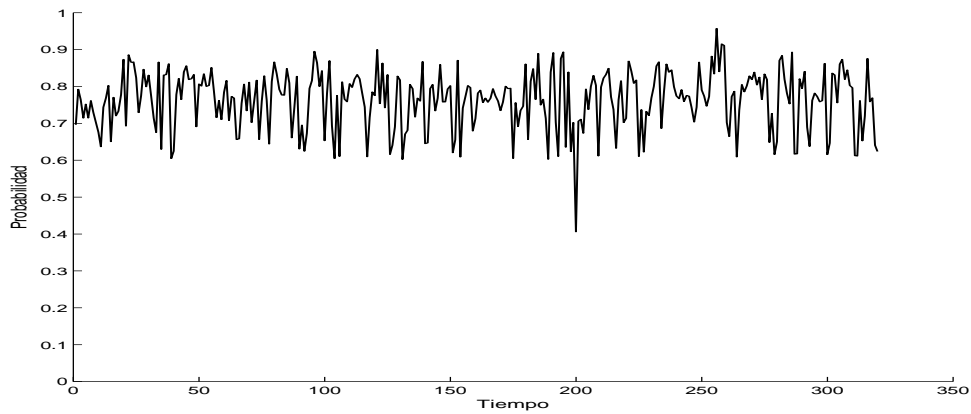
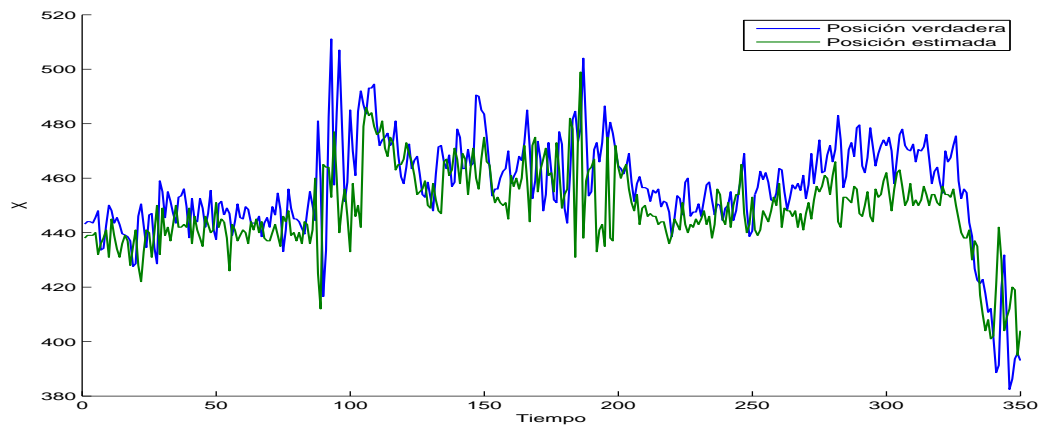
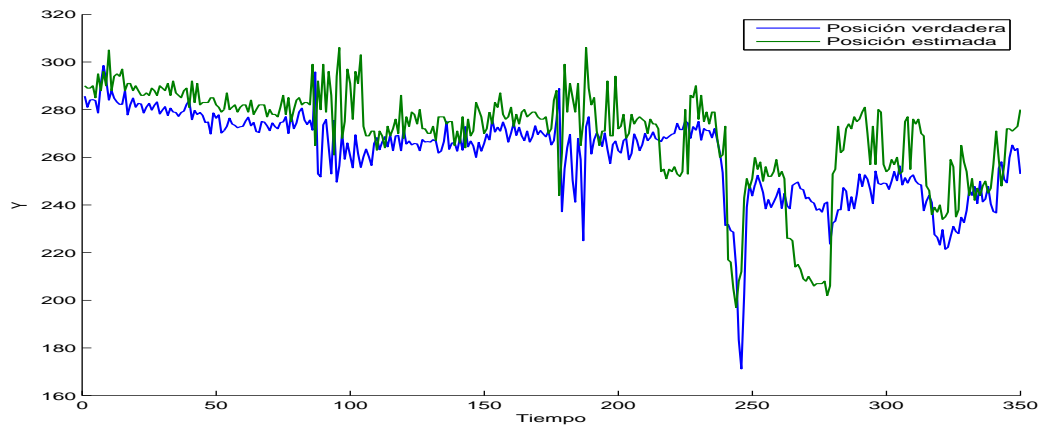


Figura 5.26: Probabilidad de las observaciones generadas con el algoritmo Condensation en el seguimiento del rostro.

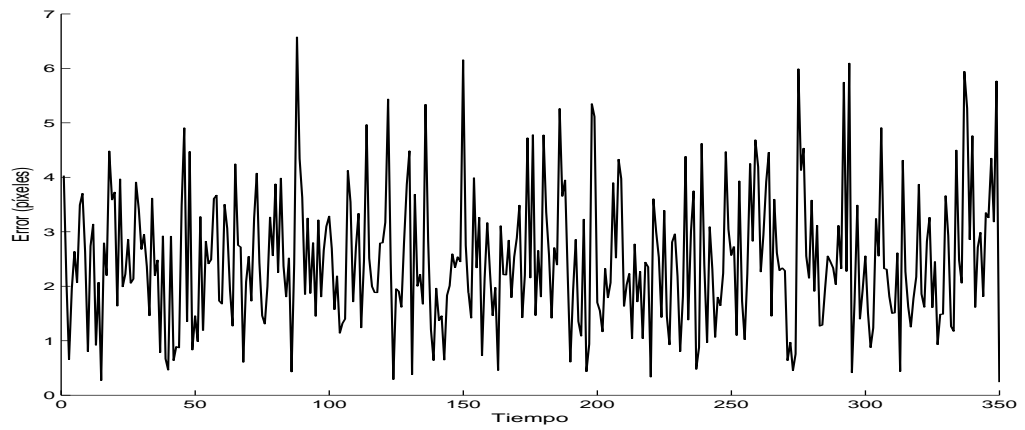
5.4 Enfoque bayesiano para el seguimiento de los ojos y el rostro



(a)



(b)



(c)

Figura 5.27: Seguimiento del rostro con el algoritmo Condensation sobre una secuencia de 350 imágenes (a) posición en la dirección x, (b) posición en la dirección y, (c) error.

5.4.2.3. Resultados del seguimiento del rostro

Los resultados a los que se ha llegado se presentan en la tabla 5.2. Los parámetros de inicialización son los mismos que en el caso del espectro visible, por lo que se recomienda revisar esta información en el capítulo anterior.

	Número de imágenes	Número de fallos	Aciertos (%)
Secuencia 1	800	40	95
Secuencia 2	646	29	95,51
Secuencia 3	600	20	96,67
Secuencia 4	324	9	97,22

Tabla 5.2: Resultados del seguimiento del rostro bajo iluminación infrarroja sobre varias secuencias de imágenes de conductores en el vehículo IvvI.

5.5. Cálculo del índice de somnolencia

Para generar el índice de somnolencia se precisa conocer el estado del ojo (abierto o cerrado) a través del tiempo y de manera continua. Este caso es similar al del espectro visible, pero el procedimiento de cálculo es distinto porque en este escenario se utilizan las características que el ojo presenta al encontrarse iluminado con la luz infrarroja.

Para conocer el estado del ojo, el sistema utiliza la información del área de la pupila detectada con el método de la sección anterior. Al cerrar el ojo, el efecto del brillo de la pupila desaparece, lo que implica que el ojo se encuentra cerrado.

En la tabla 5.3 se aprecian varios resultados generados en experimentos desarrollados sobre el vehículo IvvI y en el laboratorio. Mientras que en la figura 5.28 se observa la respuesta instantánea generada por el sistema, y en la figura 5.29 se observa la evolución de este índice sobre una secuencia de 350 imágenes tomadas en la noche sobre el vehículo IvvI.

5.5 Cálculo del índice de somnolencia

	Número de imágenes	Ojos Abiertos	Ojos Cerrados	Aciertos (%)
Secuencia 1	800	690/700	97/100	97,78
Secuencia 2	646	500/530	100/116	90,27
Secuencia 3	600	345/374	206/226	91,69
Secuencia 4	324	208/244	74/80	88,87

Tabla 5.3: Resultados del análisis del estado de los ojos bajo iluminación infrarroja sobre varias secuencias de conductores sobre el vehículo IvvI.

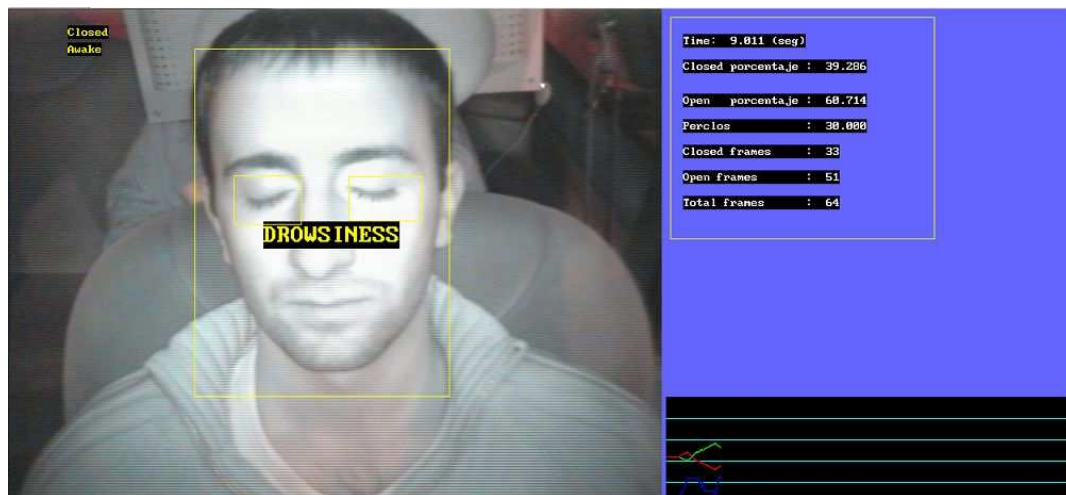


Figura 5.28: Respuesta instantánea del sistema propuesto en conducción nocturna bajo iluminación infrarroja.

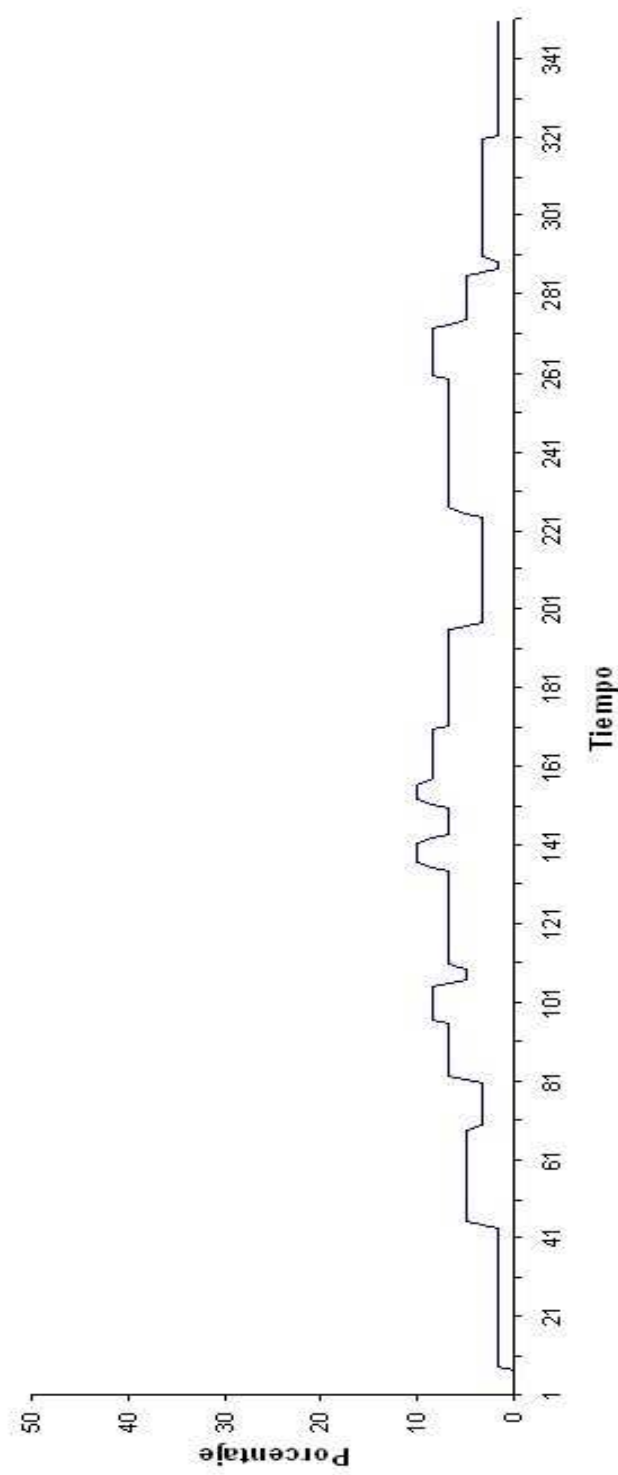


Figura 5.29: Evolución de Perclos en una secuencia de 350 imágenes.

5.6. Detección de la distracción

Al igual que en el caso del espectro visible, se plantea estimar la orientación del rostro a partir de la posición de los ojos.

5.6.1. Estimación de la orientación del rostro basada en la posición de los ojos

La orientación del rostro se calcula a través del ángulo que forman los centros de los ojos, ecuación (5.16), las restricciones para conocer si se encuentra distraído vienen dadas por los valores indicados en la ecuación (5.29). Un ejemplo sobre una secuencia compuesta de 400 imágenes se observa en la figura 5.30.

$$\left\{ \begin{array}{ll} \text{Rostro a izquierda} & \text{si } \theta > 8 \\ \text{Rostro de frente} & \text{si } |\theta| \leq 8 \\ \text{Rostro a derecha} & \text{si } \theta < -8 \end{array} \right. \quad (5.29)$$

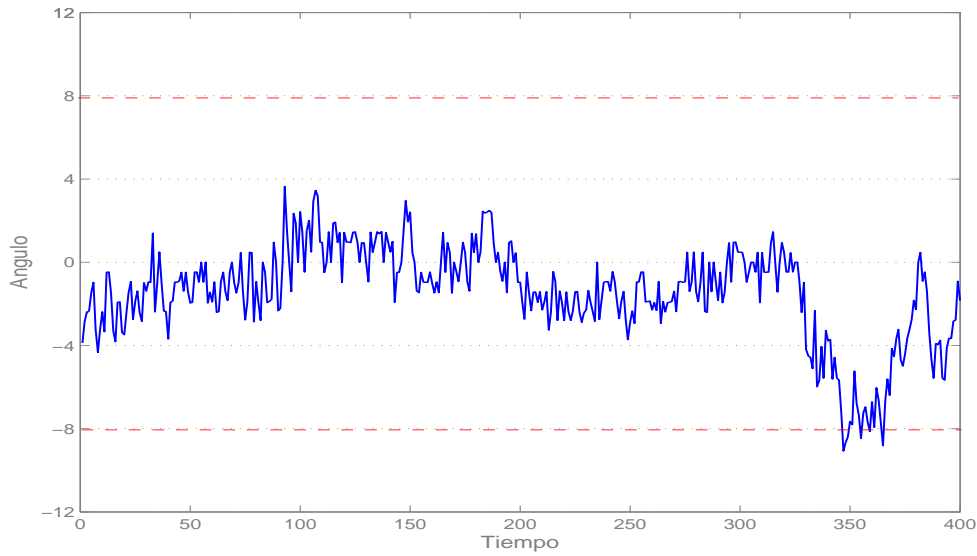


Figura 5.30: Evolución de la orientación del rostro a partir de la posición de los ojos sobre una secuencia de 400 imágenes.

5.7. Velocidad de procesamiento

En los experimentos realizados, la velocidad promedio de procesamiento fue de 275 milisegundos por fotograma, es decir, entre 3 y 4 fotogramas por segundo. Esto incluye la captura del fotograma, el procesamiento y extracción de la información y, finalmente, la visualización de los resultados por pantalla. Estos experimentos se desarrollaron en un Intel Pentium D, con 3,2 GHz. de velocidad, 2 GB. de memoria RAM y MS Windows XP.

En la tabla 5.4 se presentan los resultados del tiempo de cómputo de las pruebas realizadas.

Etapa	Tiempo
	(ms.)
Detección de los ojos	70
Detección del rostro	30
Seguimiento del rostro	110
Seguimiento de los ojos	30
Análisis del estado de los ojos	15
Análisis distracción	15
Presentación de resultados	5

Tabla 5.4: Tiempos de cómputo en cada fase del algoritmo en el infrarrojo cercano.

Conclusiones y Futuros Desarrollos

En este capítulo final se presentan dos aspectos de gran trascendencia en el desarrollo de este trabajo. El primero y el más importante se refiere a las conclusiones a las que se ha llegado al finalizar la investigación. El segundo, hace referencia a los futuros desarrollos en el campo del análisis del conductor utilizando técnicas de Visión por Computador e Inteligencia Artificial.

6.1. Conclusiones

Una vez más es necesario hacer hincapié en la investigación desarrollada en esta tesis, que se refiere al desarrollo de un sistema avanzado de ayuda a la conducción para detectar automáticamente la somnolencia y la distracción del conductor basado en información visual y técnicas de Inteligencia Artificial, para de allí partir hacia las conclusiones y las metas alcanzadas.

Así, las principales conclusiones a las que se ha llegado, en base a los resultados obtenidos, se resumen a continuación:

6.1.1. En los dos dominios

1. A partir de la información presentada al inicio de la memoria se puede concluir que debido a la excesiva cantidad de accidentes de tráfico, las autoridades de los distintos países en conjunto con varios organismos internacionales (OMS, Banco Mundial, UE, NHTSA, etc.) se encuentran trabajando con el objetivo de reducir las cifras de accidentalidad. Para ello han organizado programas de educación y prevención dirigidos a toda la población. Pero el problema es de tal magnitud que al final de la próxima década se estima que será uno de los principales causantes de muertes en el mundo. Por otra parte, la combinación de factores endógenos (carga de trabajo, edad, hora del día, etc.) y exógenos (monotonía, distracción, etc.) generan somnolencia y/o distracción; muchos de ellos son difíciles o incluso imposibles de erradicar, motivo por el cual, contribuyen con un gran porcentaje de accidentes en el sector del transporte, generando grandes pérdidas materiales y humanas. En este sentido, clasificar el tipo de accidente basado en somnolencia o distracción es complicado porque generalmente están relacionados entre sí; así, el alcohol, las drogas y las enfermedades del sueño inducen la somnolencia y/o la distracción, motivo por el cual los informes de accidentes de tráfico se encuentra solapados. Por lo tanto, existe la necesidad de investigar y desarrollar métodos y sistemas que ayuden a disminuir su impacto usando las nuevas tecnologías.
2. La complejidad del problema ha llevado a que sea absolutamente necesario dividirlo en dos partes, para así poder cubrir la conducción diurna y nocturna. En este contexto, la primera parte corresponde al algoritmo para el día (espectro visible), y la segunda, al algoritmo para la noche (infrarrojo cercano). De esta manera, el sistema construido permite analizar al conductor en situaciones reales de conducción sobre un vehículo real. Estas condiciones generan un sinnúmero de parámetros que no se pueden controlar, a pesar de ello, se puede destacar que en los experimentos realizados se han obtenido resultados satisfactorios, llegando a obtener índices fiables para medir la somnolencia y la distracción.
3. Otra aportación importante a resaltar es la versatilidad del sistema para trabajar

sobre distintos tipos de personas. En el diseño del sistema se ha utilizado características intrínsecas que no dependen de factores subjetivos como el color de los ojos o la piel, esto ha dado robustez al sistema en los dos escenarios, esto se corrobora con la información de los resultados presentados en las distintas tablas a lo largo del documento.

6.1.2. Espectro visible (conducción diurna)

1. Se ha desarrollado un nuevo método basado en la forma geométrica para localizar la posición de los ojos. Para ello se ha extraído la información necesaria para modelar dicha forma y ajustar un modelo elíptico basado en el algoritmo EM. Este método trabaja adecuadamente en ambientes con iluminación variable debido a que utiliza información invariante a los cambios de iluminación y toma en cuenta la geometría del ojo. Este método es adecuado para trabajar con personas que presenten distintos colores de la pupila, ya que es inmune a este tipo de información.
2. Se han contruido métodos eficientes para desarrollar el seguimiento del rostro y los ojos usando una estructura probabilística. Al hacer interactuar todas las capacidades de las redes neuronales artificiales con la flexibilidad del algoritmo Condensation se ha conseguido construir un efectivo método de seguimiento del rostro. Lo mismo sucede para el caso de los ojos, donde la interacción del algoritmo Condensation con la técnica de correlación de modelos ha permitido seguir eficientemente los rectángulos que contienen a los ojos, cuyos modelos se construyen una vez que se detecta el rostro, esto ha permitido tener un modelo individualizado de cada conductor, que resulta ser mucho más eficiente que tener un modelo general, construido de antemano.
3. El uso de SVM en conjunto con el filtro de Gabor ha servido para construir un método eficiente para detectar el estado del ojo. Detectar si un ojo esta abierto o cerrado es un trabajo difícil, en especial en situaciones a la interperie y mucho más cuando se tiene restringido el tiempo de cómputo bajo condiciones de tiempo real. El filtro de Gabor, al variar los parámetros de orientación y escala, resalta ciertos rasgos típicos de un ojo abierto y opaca los de un ojo cerrado, y viceversa. Por su parte SVM ha demostrado ser un método de clasificación potente en varios problemas de

Visión por Computador. Al unir las dos técnicas se ha logrado obtener un método robusto para analizar el estado de los ojos a través del tiempo. Esto ha permitido generar el índice de somnolencia PERCLOS en tiempo real y de manera automática.

4. Nuevamente, el uso de las redes neuronales artificiales ha permitido construir un sistema robusto para detectar la distracción de conductor, al analizar la orientación de su rostro a través del tiempo. Recordar que la distracción también contribuye en el aumento de las estadísticas de accidentes de tráfico en un gran porcentaje.

6.1.3. Infrarrojo cercano (conducción nocturna)

1. Para abordar el problema de la somnolencia en conducción nocturna se ha construido un complejo sistema de visión basado en la luz infrarroja del tipo cercano. Este sistema es capaz de separar la señal de video en dos campos, denominados par e impar, que son la base para localizar la posición de los ojos. Lo novedoso de su diseño, radica en que es posible controlar la cantidad de luz radiada mediante dos maneras: (i) variando la corriente de los leds exteriores e interiores, y (ii) modificando la distancia de los leds exteriores respecto al anillo central de leds. Estas incorporaciones han sido beneficiosas para ajustar la calidad de las imágenes obtenidas. Además, se ha logrado construir un sistema de visión de tamaño reducido, ideal para distintas aplicaciones del tipo análisis del conductor. Por lo tanto, el sistema ha sido instalado en el vehículo experimental IvvI sin ninguna dificultad y sin generar molestias ni inconvenientes al conductor.
2. En cuanto al desarrollo del algoritmo, se ha presentado una metodología para detectar los ojos basada en la imagen diferencia, su imagen de bordes y los algoritmos de umbralización desarrollados para segmentar entre el fondo (colores oscuros) y los potenciales ojos (colores blancos). Una vez obtenido el conjunto de cuerpos se los filtra bajo dos estrategias, la primera consiste en eliminar los cuerpos que no cumplan ciertas restricciones geométricas, y la segunda consiste en discriminar utilizando el clasificador SVM entrenado con imágenes de ojos abiertos y no-ojos, procesadas con el filtro de Gabor. Esta doble estrategia ha permitido localizar la posición de las pupilas y por ende la de los ojos. Con esto se ha conseguido que el algoritmo sea

robusto bajo condiciones extremas para detectar la posición de los ojos abiertos y dar inicio al algoritmo de seguimiento.

3. Se ha implementado un método robusto de seguimiento de los ojos basado en el algoritmo Condensation y el brillo de las pupilas bajo la iluminación infrarroja. Para la evaluación de la función de densidad de las observaciones se ha recurrido a una doble estrategia que consiste en construir una función de densidad de probabilidad triangular que toma en consideración el nivel de gris de la imagen diferencia favoreciendo a los píxeles de mayor valor, luego, se ha tomado en consideración la posición de los cuerpos generados con el método de detección de los ojos.
4. La doble estrategia desarrollada para localizar el rostro permite localizarlo de manera eficaz y sencilla. La primera estrategia consiste en utilizar la función de varianza vertical, una vez modificado el contraste de la imagen con infrarrojos, para localizar el cuerpo que corresponde al conductor. La segunda estrategia utilizada ha consistido en utilizar la posición de los ojos para construir un modelo del rostro; para comprobar que la zona estimada se trata efectivamente de uno se ha utilizado una red neuronal, entrenada con imágenes de rostros y no-rostros bajo iluminación infrarroja y procesadas con el filtro de Gabor. La estrategia utilizada ha permitido corregir las deficiencias que se pueden presentar bajo iluminación infrarroja.
5. Para el seguimiento del rostro, nuevamente se ha recurrido al uso del algoritmo Condensation, donde el modelo dinámico y la estrategia de seguir han sido las mismas que en el caso del espectro visible.
6. Para implementar el índice PERCLOS, que mide el estado de somnolencia del conductor, se ha utilizado la información del área de la pupila sobre el tiempo. Es decir, el estado del ojo está determinado exclusivamente por el efecto que produce la luz infrarroja sobre la pupila. Para conocer si el ojo se encuentra abierto se verifica si existe algún cuerpo en la zona predicha sobre la imagen binarizada. Cuando el seguimiento sobre los ojos falla se utiliza la posición estimada del rostro para verificar el estado del ojo.

6.2. Futuros desarrollos

Si bien queda claro que el trabajo desarrollado es bastante satisfactorio para los objetivos planteados y las metas alcanzadas durante distintos experimentos desarrollados sobre el vehículo IvvI, no es menos cierto que aún resta mucho trabajo por hacer hasta llegar a construir un sistema que sea capaz de monitorizar todas las actividades del conductor y que sea 100 % eficaz, robusto y fiable para ser instalado en un vehículo comercial. Es por ello que existen diferentes líneas de trabajo de gran interés en las cuales se puede seguir investigando y, por ende, mejorando el algoritmo propuesto. Entre los principales trabajos que se pueden desarrollar como continuación de ésta tesis son:

1. Incluir la información proveniente del exterior, por ejemplo: la temperatura, la hora del día, las condiciones del clima, etc., para mejorar el índice de somnolencia e interactuar de mejor manera con la información medida sobre el conductor.
2. Para avanzar en el estudio de la somnolencia y/o la distracción del conductor es inevitable ir un paso más allá de la construcción de este sistema u otro relacionado, para ello, es necesario generar información utilizando conductores profesionales en condiciones reales de conducción. Con esta información se podrán contruir *patrones de conducción*, con lo cual se conseguirá desarrollar estudios mucho más minuciosos del comportamiento del conductor en su ambiente laboral, logrando predecir su estado.
3. Incluir un módulo para analizar las expresiones faciales del conductor, en este punto, se podrían utilizar los modelos AAM (active apparence models) o en los modelos ASM (active shape models) propuestos por Cootes *et al.* [49], [136]. Las expresiones faciales cambian de acuerdo al estado del conductor [60], por ejemplo, se podría conocer si se encuentra contento, cansado, triste, etc. Este módulo permitiría disponer de información que vaya un paso más alla del análisis de la somnolencia [125], es decir, se podrían generar nuevos índices que cuantifiquen el comportamiento humano.

Apéndice

Plataforma experimental IvvI

IvvI es una plataforma experimental diseñada para el desarrollo de sistemas de ayuda a la conducción basados en técnicas de Visión por Computador [90]. En la figura A.1 se aprecia la imagen del vehículo marca Renault del tipo *Twingo* que alberga los sistemas desarrollados durante la elaboración de varias tesis doctorales y que han generado este prototipo de la ingeniería y la tecnología, que pretende constituirse en un copiloto que permita ayudar a disminuir los errores (voluntarios o involuntarios) generados durante el proceso de conducción.

Para comenzar a describir los sistemas que componen IvvI es necesario observar la figura A.2 donde se aprecia como están interconectados los sistemas y cual será su función dentro del vehículo. Los módulos que están soportados actualmente se dividen en seis partes: detección de señales de tráfico, detección de vehículos, detección de peatones, detección de las líneas de la carretera, y finalmente, detección de la somnolencia del conductor. Estos módulos serán analizados con detalle más adelante en este capítulo; pero antes se describen brevemente los distintos componentes hardware utilizados como apoyo para desarrollar los distintos módulos.

IvvI está equipado con un convertidor *DC/AC* conectado a una batería; a través del cual se suministra la corriente eléctrica necesaria para que operen los ordenadores, las cámaras y el GPS. Existen dos ordenadores en la parte posterior del vehículo, ver la figura A.3 izquierda, que son los encargados de procesar la información proveniente de las cámaras, además, se incluye una pantalla, un ratón y un teclado. En la parte frontal derecha se localiza un sistema compuesto por dos cámaras en blanco y negro, que forman



Figura A.1: Plataforma experimental Ivvi en una presentación desarrollada en el IFEMA durante la Semana de la Ciencia de Madrid.

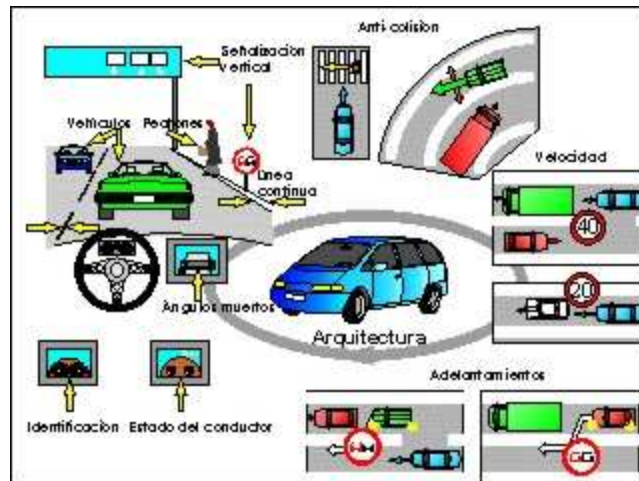


Figura A.2: Arquitectura de los sistemas que componen la plataforma Ivvi.

un sistema de visión estéreo para detectar las líneas de la carretera y los peatones en el espectro visible, y una cámara a color para detectar las señales de tráfico, ver la figura A.3 (b).



Figura A.3: Interior de Ivvi: (a) sistema de procesamiento y pantalla, (b) cámaras para la detección de peatones, señales de tráfico y carriles de la carretera.

En la parte frontal izquierda, sobre el volante, se localiza una cámara a color para capturar imágenes del conductor, figura A.4. Mientras en la parte exterior izquierda, sobre el espejo, se ubica una cámara térmica para detectar peatones en la noche, figura A.5.

En las próximas secciones se describirá detalladamente los sistemas que han sido anota-



Figura A.4: Camaras Interiores de Ivvi: Cámara con iluminación infrarroja colocada en el panel del espejo superior para captar información del conductor durante la noche.

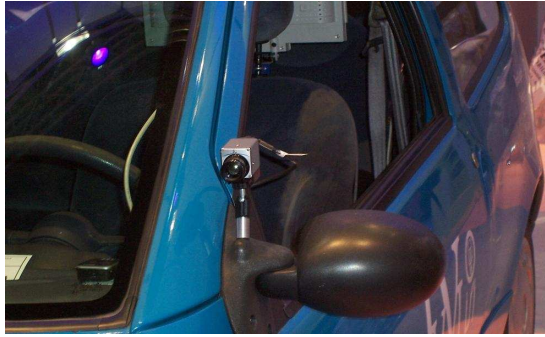


Figura A.5: Exterior Ivvi: Cámara térmica colocada en el soporte del espejo exterior para captar información de los peatones.

dos y como intervienen cada una de las cámaras que forman el sistema de percepción de Ivvi.

A.1. Sistema de detección de vehículos

La detección de vehículos es fundamental para las siguientes tareas: *seguimiento en peatón*, los vehículos circulan en grupo a altas velocidades y con distancias de separación pequeñas; *stop and go*, básicamente es lo mismo que el problema anterior pero para el caso de conducción dentro de una ciudad; *ángulo muerto*, el sistema debe detectar que existe otro vehículo aproximándose lo que impedirá que nuestro coche inicie el adelantamiento; *supervisor de maniobras* propias y las de los demás vehículos [72].

A.2 Sistema de detección del carril

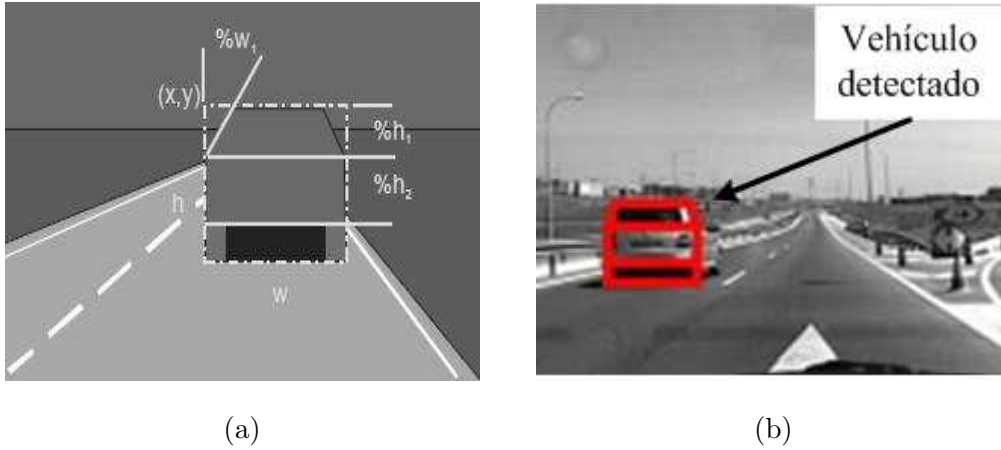


Figura A.6: Detección de vehículos utilizando modelos geométricos, (a) modelo geométrico que describe la forma del vehículo, (b) resultado de la detección del vehículo.

En la figura A.6 se aprecia como el sistema es capaz de detectar la posición del vehículo que se encuentra al frente sobre la carretera.

A.2. Sistema de detección del carril

Conocido también con el nombre de *sistema de detección ante el alejamiento involuntario del carril*. Este sistema trabaja con las imágenes tomadas por las cámaras instaladas en la parte frontal del vehículo, es decir, capturas por el sistema estéreo que se presentó en la figura A.3 (b).

El sistema está encargado en determinar la posición del vehículo respecto a las líneas que marcan el límite del carril; para ello, detecta las líneas de la carretera [74] y las clasifica en líneas continuas o semicontinuas, utilizando la transformada de Fourier [73], el sistema es robusto gracias a que incorpora un modelo de la carretera [71]. Una alarma se dispara cuando la trayectoria del vehículo se aleja de la ideal, que es mantenerse dentro del carril establecido por la autoridad de tráfico.

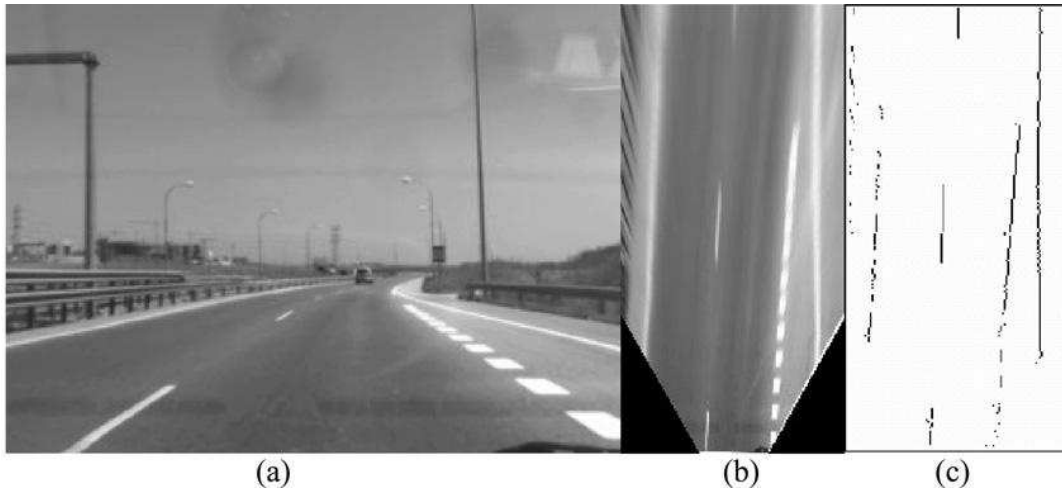


Figura A.7: Detección del carril de la carretera, (a) imagen de la carretera, (b) vista aérea, (c) segmentación de las líneas de la carretera.

A.3. Sistema de detección de peatones

Los peatones están expuestos a sufrir accidentes y son los más vulnerables e indefensos en la circulación, por el hecho que los vehículos y los peatones actúan en entornos exteriores.

Los accidentes de tráfico donde se ven involucrados los peatones son bastante frecuentes [148]. Muchos de ellos ocurren en cruces peatonales donde el peatón tiene el privilegio para cruzar la calle, pero, donde a la vez, el conductor no presta la suficiente atención, generando un accidente donde el más perjudicado es el peatón. Según varios reportes de accidentes de tráfico se estima que el 20 % de ellos involucran a un peatón [14].

Actualmente IvvI alberga dos módulos que detectan peatones, uno en el día y otro en la noche. El del día utiliza la información de simetría calculada sobre la imagen de disparidades [15], [17], [16] capturada por el sistema de visión estéreo. El de la noche utiliza la información térmica para localizar los *cuerpos calientes* dentro de la imagen, y luego la información de la simetría permite filtrar los cuerpos que no son peatones hasta obtener los que si corresponden.

En la figura A.8 se aprecia un ejemplo del sistema durante el día, mientras que en la

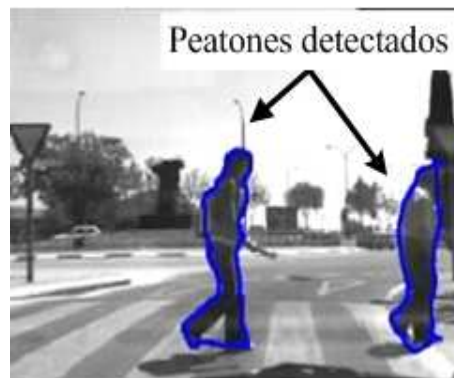


Figura A.8: Detección de peatones en el día.



Figura A.9: Detección de peatones en la noche, las zonas de color blanco corresponden a cuerpos calientes.

figura A.9 se aprecia un ejemplo del sistema descrito para la noche.

A.4. Sistema de detección de señales de tráfico

Las primeras señales de tráfico modernas, específicamente la señal de PARE, se diseñó en Alemania en el año 1.892. Esta consistía en una calavera metálica que se iluminaba en la noche. La implantación de esta señal provocó pánico entre los viajeros de la época. En la actualidad, las señales de tráfico son de gran utilidad para ayudar a ordenar el tráfico, porque su finalidad es señalar a los conductores un acatamiento o una prohibición de tráfico.

En los últimos años, la detección de señales de tráfico ha recibido un creciente interés

motivado por las siguientes razones:

- Mantenimiento de autopistas: permitiendo con esto la automatización de la tediosa labor de determinar si las señales se encuentran en buen estado y gozan de buena visibilidad, trabajo que actualmente lo hace un operario.
- Inventario de señales en ciudades: En este entorno las señales no están siempre visibles.
- Sistemas de ayuda a la conducción: su interpretación facilita supervisar la velocidad y la trayectoria del vehículo.

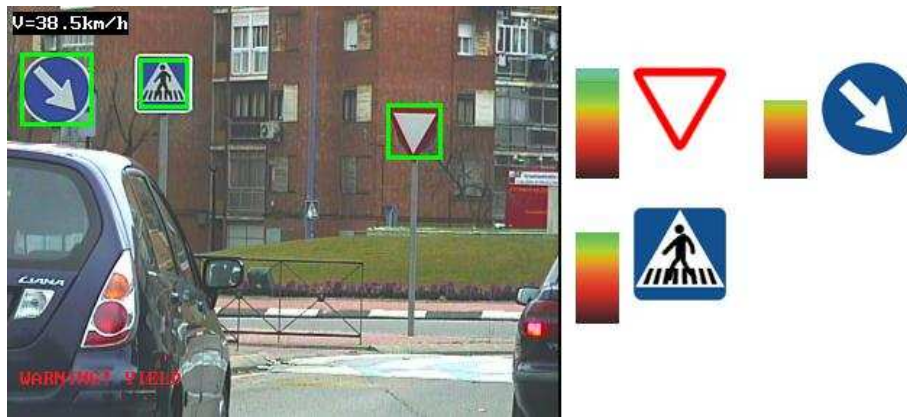


Figura A.10: Detección de las señales de tráfico.

El módulo de detección de señales de tráfico utiliza la información del color y de la forma para detectar las señales y la correlación normalizada para clasificarlas [113], [33], [34], [32], [31], [29], [30], [114]. En la figura A.10 se observa un ejemplo del sistema que actualmente se encuentra instalado en Ivvi.

A.5. Sistema de detección de la somnolencia

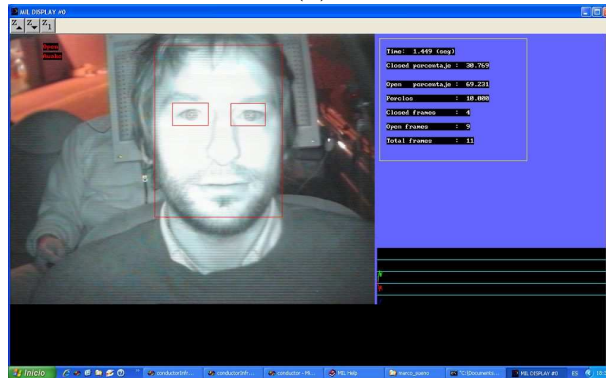
La somnolencia es un mal que afecta, en mayor o menor medida, a todos los conductores, y es la causante de un importante porcentaje de los accidentes de tráfico. La construcción de un sistema para detectar la somnolencia ha sido el principal motivo que ha dado lugar al desarrollo de esta tesis doctoral y que ha sido descrito en detalle a lo

A.5 Sistema de detección de la somnolencia

largo de los capítulos que componen el presente documento. Es así, que se ha elaborado, principalmente, el software que se presenta en la figura A.11.



(a)



(b)

Figura A.11: Sistema de detección de la somnolencia, (a) en el espectro visible, (b) en el infrarrojo cercano.

Diseño y Construcción del Sistema de Percepción

En el capítulo tres se introdujo la necesidad de construir el hardware que forma parte del sistema de detección de la somnolencia. Por lo tanto, es aquí donde se expone y desarrolla la electrónica asociada para la construcción del sistema de iluminación y su integración con el sistema de visión. También se presentan las pruebas desarrolladas para garantizar la seguridad del conductor, en especial de sus ojos al ser irradiados con la luz infrarroja.

B.1. Sistema de visión en el infrarrojo cercano

Con los antecedentes presentados en el capítulo cuatro se procede a construir el sistema de visión basado en iluminación infrarroja del tipo cercano. Recordar que los sistemas que se diseñan son los siguientes:

- Sistema de visión.
- Sistema de iluminación.
- Sistema de sincronización.

B.1.1. Elementos del sistema de visión

Los componentes que forman parte del sistema de visión con iluminación infrarroja son:

-
- Una cámara GC400 CP.
 - Un circuito integrado LM1881 y dos circuitos integrados de LM324.
 - Un juego de transistores de amplificación BD437.
 - Un juego de diodos leds¹ SFH 487 y ELJ-810-248 B.
 - La tarjeta de adquisición de datos Morphis.
 - Varias resistencias, tres potenciómetros, cables de conexión y una caja.

B.1.2. Sistema de visión

El sistema de visión esta compuesto esencialmente por la cámara GC400 CP que es una cámara a color que proporciona una imagen analógica en video compuesto (PAL), de la cual se obtiene la señal de sincronismo que será la entrada del circuito integrado LM1881.



Figura B.1: Cámara GC400CP.

Por otra parte, la mayoría de los sensores CCD que incluyen las cámaras son sensibles en el espectro visible y a su vez en el infrarrojo cercano, por tanto la imagen tomada capta este tipo de radiación. Para evitar ese efecto se suele colocar un filtro óptico que elimina el espectro IR. En este caso no se quiere eliminarlo y por lo tanto se quita el filtro. En

¹acrónimo del inglés Light Emitting Diode

B.1 Sistema de visión en el infrarrojo cercano

imagen sensor	sharp 1/4-inch interline colour CCD
horizontal resolution	350 TVL colour - 512x582 pixels
electronic shutter	1/50 - 1/1000.000 seconds
sensitivity	1.0 lux & f1.2
signal to noise ratio	greater than 45dB (agc off)
white balance	automatic
lens	3.6 mm. conical pinhole
power consumption	100 mA.
video output	1Vp - p 75 Ω
dimensions	32mm x 32mm
power supply	12V \pm 10%
operating temperature	-10° 50°
operating humidity	10 % 90 %

Tabla B.1: Características de la cámara GC400CP.

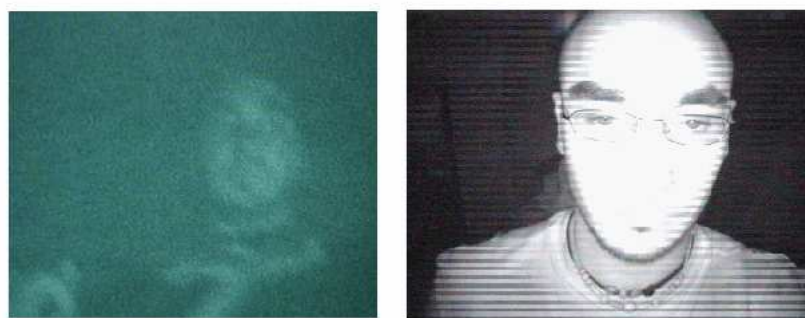
este modelo de cámara se encuentra acoplado a la lente cónica, por lo que se procede a quitarlo de la lente, como indica la figura B.2.



Figura B.2: Filtro IR del sistema de visión.

Para apreciar el efecto de eliminar el filtro se procede a tomar dos imágenes con una iluminación IR en un entorno totalmente a oscuras.

Finalmente, por las dimensiones de la cámara que son bastante reducidas, presenta un



(a)

(b)

Figura B.3: Imágenes tomadas con y sin filtro IR, (a) con filtro IR, (b) sin filtro.

tamaño adecuado para ubicarla en lugares donde no interfiera la visión y comodidad del conductor.

B.1.3. Sistema de iluminación

En esta parte se describen los elementos radiantes y el diseño del circuito que los controla.

B.1.3.1. Diseño

La cámara proporciona una imagen en formato PAL. En este sistema se llama cuadro a una imagen completa. Esta imagen o cuadro se compone de una serie de líneas horizontales que son barridas totalmente cada 50 Hz. De estos 50 Hz la mitad son para las líneas de campo pares y la otra para los impares. Pues bien, la idea es proporcionar una imagen donde las líneas pares e impares sean iluminadas con diferentes intensidades de luz y ángulos de incidencia en combinación para conseguir el efecto deseado de iluminación sobre el ojo, es decir, resaltar el brillo de la pupila, y poder realizar posteriormente un procesamiento digital de la imagen.

Para tal fin se necesita contar con dos sistemas radiantes de infrarrojos que se iluminen de acuerdo a las necesidades requeridas.

La diferencia de intensidad lumínica que capta la cámara se origina por el ángulo de

B.1 Sistema de visión en el infrarrojo cercano

incidencia de la luz, de forma que dependiendo del ángulo la cantidad de luz reflejada será diferente. Esto se conoce como la ley de Snell:

Considérense dos medios caracterizados por los índices de refracción η_1 y η_2 separados por una superficie S , figura B.4. Los rayos de luz que atraviesen los dos medios se refractarán en la superficie variando su dirección de propagación dependiendo del ratio entre los índices de refracción η_1 y η_2 .

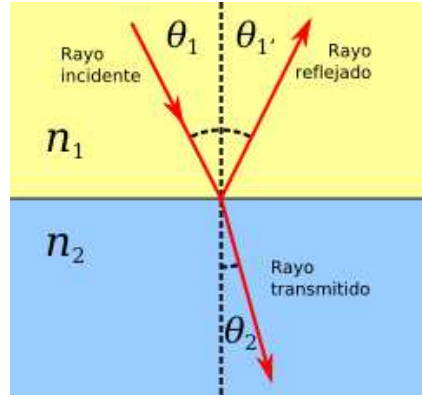


Figura B.4: Ley de Snell.

Para un rayo luminoso con un ángulo de incidencia θ_1 sobre el primer medio, ángulo entre la normal a la superficie y la dirección de propagación del rayo, se tiene que el rayo se propaga en el segundo medio con un ángulo de refracción θ_2 cuyo valor se obtiene por medio de la ley de Snell

$$\eta_1 \sin(\theta_1) = \eta_2 \sin(\theta_2) \quad (\text{B.1})$$

Obsérvese para el caso $\theta_1 = 0$ (rayos incidentes de forma perpendicular a la superficie) los rayos refractados emergen con un ángulo $\theta_2 = 0$ para cualquier $\eta_1 = \eta_2$.

La simetría de la ley de Snell implica que las trayectorias de los rayos de luz son reversibles. Es decir, si un rayo incidente sobre la superficie de separación con un ángulo de incidencia θ_1 se refracta sobre el medio con un ángulo de refracción θ_2 , entonces un rayo incidente en la dirección opuesta desde el medio 2 con un ángulo de incidencia θ_2 se refracta sobre el medio 1 con un ángulo θ_1 .

Una regla cualitativa para determinar la dirección de la refracción es que el rayo en el medio de mayor índice de refracción se acerca siempre a la dirección de la normal a la superficie. La velocidad de la luz en el medio de mayor índice de refracción es siempre menor.

La ley de Snell se puede derivar a partir del principio de Fermat, que indica que la trayectoria de la luz es aquella en la que los rayos de luz necesitan menos tiempo para ir de un punto a otro. En una analogía clásica propuesta por el físico Richard Feynman, el área de un índice de refracción más bajo es substituida por una playa, el área de un índice de refracción más alto por el mar, y la manera más rápida para un socorrista en la playa de rescatar a una persona que se ahoga en el mar es recorrer su camino hasta ésta a través de una trayectoria que verifique la ley de Snell, es decir, recorriendo mayor espacio por el medio más rápido y menor en el medio más lento girando su trayectoria en la intersección entre ambos.

Por lo tanto, queda claro que la disposición de los sistemas radiantes es fundamental para poder obtener la funcionalidad que se busca. En este punto se puede observar soluciones propuestas en otros diseños que suelen estar caracterizadas por la disposición de dos anillos concéntricos alrededor de la cámara de forma que su iluminación sea alternante.

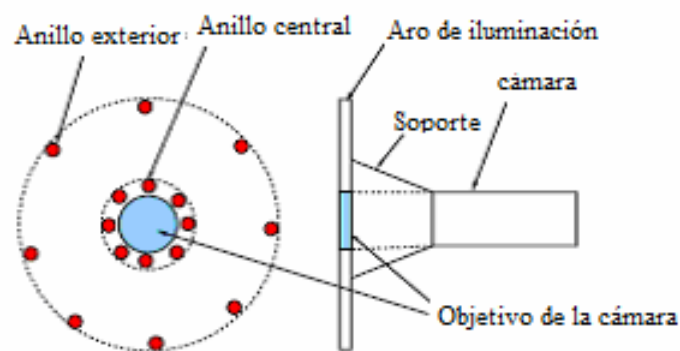


Figura B.5: Diseño genérico del sistema de visión con iluminación infrarroja.

B.1 Sistema de visión en el infrarrojo cercano

B.1.3.2. Emisores IR y electrónica asociada

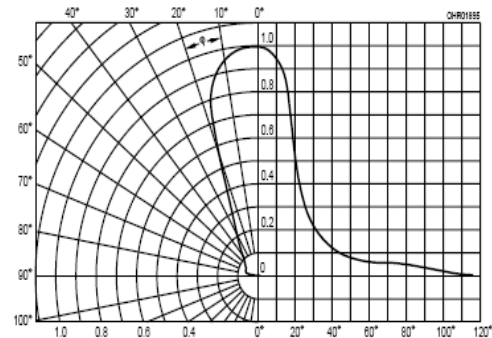
La radiación emitida por un LED está limitada por la física del semiconductor y la estructura que lo componen, siendo la eficiencia cuántica de un LED en torno al 20 % y 40 % normalmente.

Se elegirán dos tipos diferentes de emisores IR. Un grupo será el que constituye el anillo central alrededor de la cámara, mientras que el segundo grupo constituye la iluminación exterior y se encuentra en los dos extremos de la cámara.

Los leds del anillo interior son los del modelo SFH 487 de Osram, figura B.6 (a). Este modelo es un emisor infrarrojo GaAlAs en la longitud de onda 880 nm., la respuesta de intensidad de radiación es de alrededor de 12mW/sr ($IF = 100$ mA, $tp = 20$ ms), soportando una corriente máxima de 100 mA y disipando hasta 200 mW. A la hora de emitir tienen una emisión máxima de 25 mW ($IF = 100$ mA, $tp = 20$ ms.) y un diagrama de radiación como el de la figura B.6 (b).



(a)



(b)

Figura B.6: Diagrama de radiación del SFH 487, (a) Imagen del led SFH 487 de Osram y (b) su diagrama de radiación.

Por lo tanto con este diseño se obtendrá un anillo central dirigido con una abertura de alrededor de unos 20°.

Para el segundo grupo de elementos emisores se debe buscar una solución que sustituya al diseño de un segundo anillo por un par de emisores IR y aun así no perder efecto ni

calidad en la iluminación. En vez de construir un grupo de emisores se ha decidido utilizar el denominado Jumbo. Este elemento no es más que un emisor ELJ-810-248 B de AlGaAs que fabrica la casa EPIGAP Optoelektronik GmbH.

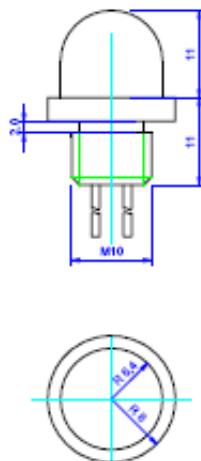


Figura B.7: Led tipo jumbo ELJ-810-248 B.

Este elemento presenta hasta una disipación de 3 W y una corriente de hasta 250 mA, produciendo una potencia de radiación de hasta 175 mW y una intensidad de radiación de alrededor de 400 mW/Sr. Como se puede ver este único elemento equivale a alrededor de unos 6 emisores SFH-487, por lo que se gana mayor versatilidad en cuanto a tamaño. Igualmente será el elemento que más radie y por tanto el que mayor control necesite. Se utilizan dos jumbos colocados a ambos extremos de la cámara, equidistantes.

En total se utilizan 10 emisores SFH 487 para formar el anillo central y dos emisores ELJ-810-248-B que para los extremos. No obstante destacar que para el estudio experimental los ELJ-810-248 B se desemplazarán sobre un carril a ambos extremos de forma que se podrá graduar la separación de estos respecto a la posición de la cámara.

Para el control y alimentación de los leds se propone el circuito de la figura B.8 (a).

Este circuito está compuesto por un transistor BD437 diseñado para soportar potencias de disipación medias, puede llegar a soportar hasta $I_c = 2A$ lo que viene bien porque no generará problemas para alimentar los leds. La figura B.8 (b) muestra este transistor.

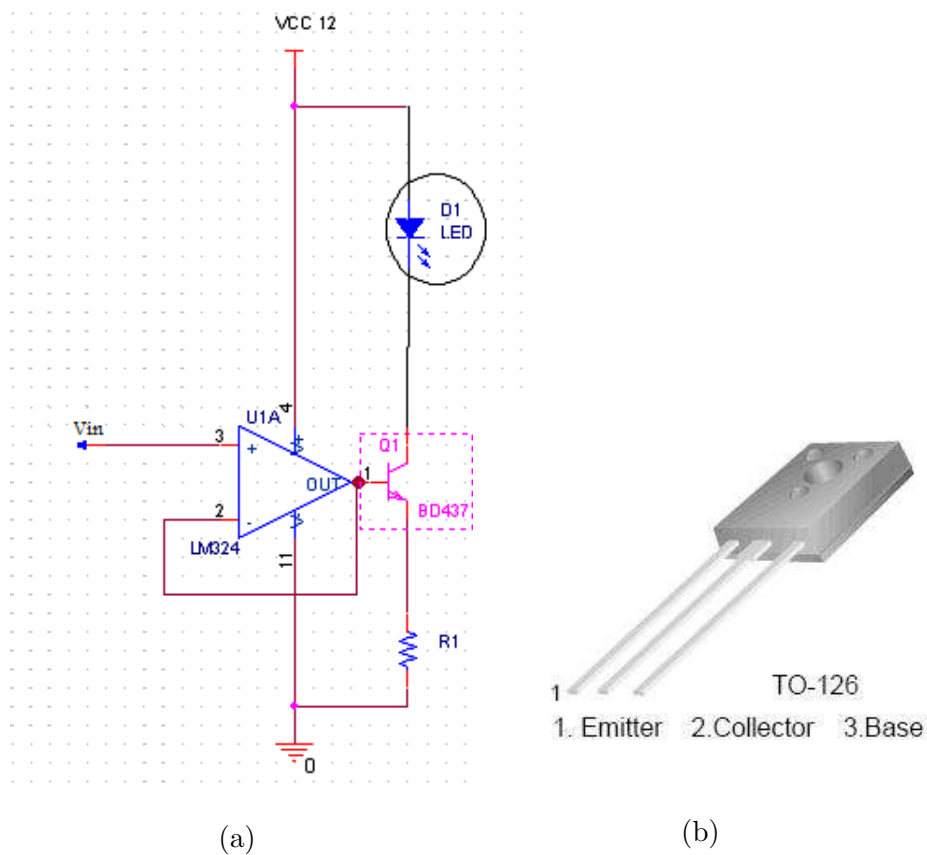


Figura B.8: (a) Fuente de alimentación genérica de los elementos emisores, (b) transistor BD437.

Aunque aún no se ha especificado detalladamente el funcionamiento del sistema completo, los elementos sensores irán conmutando entre ON/OFF cada 25 Hz. Las características de este transistor indican que funciona sin ningún problema hasta 3 MHz y por tanto no supondrá un factor limitante del diseño.

El segundo elemento con el que se logra estabilizar la corriente y además aumentar la impedancia de entrada frente a V_{in} es un amplificador operacional LM324. Se utiliza una configuración de buffer en realimentación negativa con ganancia unidad. Mediante V_{in} se controla la tensión de base y por tanto la corriente que se proporciona al elemento emisor IR. Igualmente la respuesta en frecuencia de este elemento está dentro de lo que se requiere.

Para el valor de $R1$ se selecciona una resistencia de 33Ω y una potencia adecuada para garantizar que pueda soportar los posibles calentamientos.

Aunque más adelante se destacará las especificaciones técnicas del diseño, enunciar que V_{in} será una señal cuadrada que proporcionará el sistema de sincronización y tendrá aproximadamente 7,5 Vpp.

Una vez presentado el circuito genérico se pasa a decidir que configuración de montaje se tendrá con los leds que forman el anillo central. Como el integrado de LM324 tiene 4 amplificadores operacionales diferentes, se ha optado por utilizar tres fuentes de corriente diferentes, de forma que dos controlarán el anillo central, con cinco leds cada una, mientras que la tercera controlará a los dos jumbos.

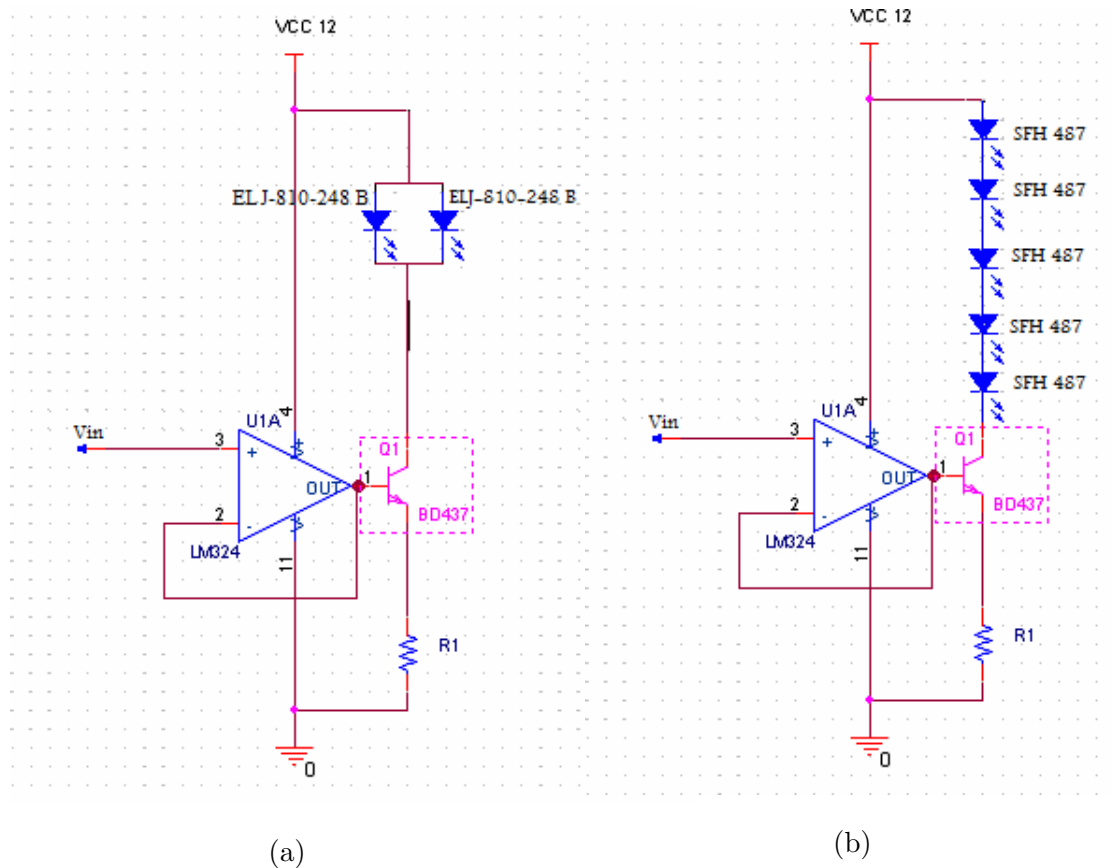


Figura B.9: (a) Fuente de alimentación de los jumbos, (b) fuente de alimentación de los leds SFH 487. En total el anillo central se compone de 10 leds y por lo tanto de 2 fuentes de alimentación.

B.1.4. Sistema de sincronización

Como se adelanto al principio de este capítulo, para realizar el futuro procesado de la imagen lo que se desea es obtener en cada captura una imagen donde las líneas pares se encuentren iluminadas con un grupo de emisores de luz y que las líneas impares se encuentren iluminadas con otro grupo [124], [78], [129].

Por tanto este objetivo será conseguido con un diseño que sincronice la señal que se obtiene de la cámara del sistema de visión con la correspondiente entrada V_{in} para cada grupo de elementos emisores del sistema de iluminación.

B.1.4.1. Sistema de video compuesto PAL

La señal de video compuesto no contiene información de sonido (ésta ha de viajar a través de otro cable), pero contiene todo lo necesario para generar la imagen que aparecerá por pantalla, esto es, pulsos de sincronismo, señal de luminancia y crominancia. Los pulsos de sincronismo, o simplemente sincronismos, son las señales que hacen que el movimiento de haz catódico por la pantalla esté sincronizado con la señal de video. Existen sincronismos verticales, que indican el comienzo de un nuevo campo y hacen que el haz se sitúe en la esquina superior izquierda de la pantalla, y sincronismos horizontales, que indican el inicio de una nueva línea y hace que el rayo catódico se sitúe al comienzo de ésta. Entre cada dos sincronismos horizontales se encuentra la señal de luminancia. Cuanto más alta es la tensión de esta señal, más brillará el rayo catódico. La información que indica el tono y la saturación del color se encuentra modulada sobre una onda portadora de alta frecuencia en la propia señal de luminancia.

En la figura B.10 aparecen representadas dos líneas de la imagen separadas por un sincronismo horizontal. Los sincronismos aparecen como pulsos negativos respecto al nivel de negro de la imagen. Este nivel es la tensión mínima que puede tener la señal de luminancia, que en la figura aparece como una línea horizontal discontinua.

En los sistemas PAL², la información de crominancia se introduce modulando en cuadratura una subportadora “de color”. La frecuencia de la subportadora es de 50 Hz en

²acrónimo del inglés Phase Alternating Line



Figura B.10: Señal de video compuesto vista desde el osciloscopio.

el PAL (Salvo el PAL de 60 Hz de Brasil). Entonces los múltiplos (armónicos) de estas frecuencias deben respetarse para mantener separadas la luminancia y crominancia. El nombre lo recibe del modo en que la información de crominancia (color) es transmitida, invirtiéndose en cada línea de forma que los posibles errores en fase se cancelen entre sí y de esta forma poder subsanarse de forma automática. De todas las especificaciones que el ITU-R BT.470-6 ofrece, las más importantes serán las siguientes:

- Scan Lines: 625
- Field rate (frames per second): 25
- Horizontal frequency: 15,625 Hz.
- Vertical frequency: 50 Hz.

De estas especificaciones se observa que el barrido vertical completo se realiza cada 50 Hz., de los cuales 25 son invertidos para el barrido de un campo (par) y los otros 25 para los del campo restante (impar).

B.1 Sistema de visión en el infrarrojo cercano

B.1.4.2. Decodificación de las señales de campo

Se necesita obtener las señales de sincronismo de los campos par e impar de la señal de video compuesto. Para tal fin, se utiliza el circuito integrado LM1881 que es un separador de señales de sincronismo encargado de extraer el campo par e impar de las imágenes capturadas por la cámara.

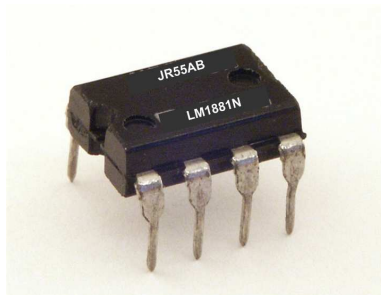


Figura B.11: Integrado LM1881N.

Este integrado es un separador de señales de sincronismo que obtiene eventos temporales de la señal de video compuesto, como pueden ser la señal de sincronismo vertical, señal de “burst”, señal de campo par e impar, etc. [137]. El esquema correspondiente al integrado es el de la figura B.12.

Del circuito integrado se pueden obtener diferentes señales de las cuales para el objetivo de este trabajo se usará exclusivamente la señal que proporciona el pin 7 (Odd/Even Output). La señal que produce el Pin 7 del campo Odd/Even es una señal que esta a nivel alto para el caso de campo par y nivel bajo para el caso de campo impar.

El diseño de control del sistema de iluminación esta planteado para funcionar siempre y cuando la señal sea positiva y con valores mayores de 0,6 V (Nivel Vbe del TRT correspondiente). Por este motivo se divide la señal obtenida del PIN 7 en dos señales donde una estará conectada directamente a un sistema de iluminación, controlando su encendido, y la otra pasará por un inversor cuya salida se conectará al otro sistema de iluminación y así conseguir la alternancia de las señales.

Connection Diagram

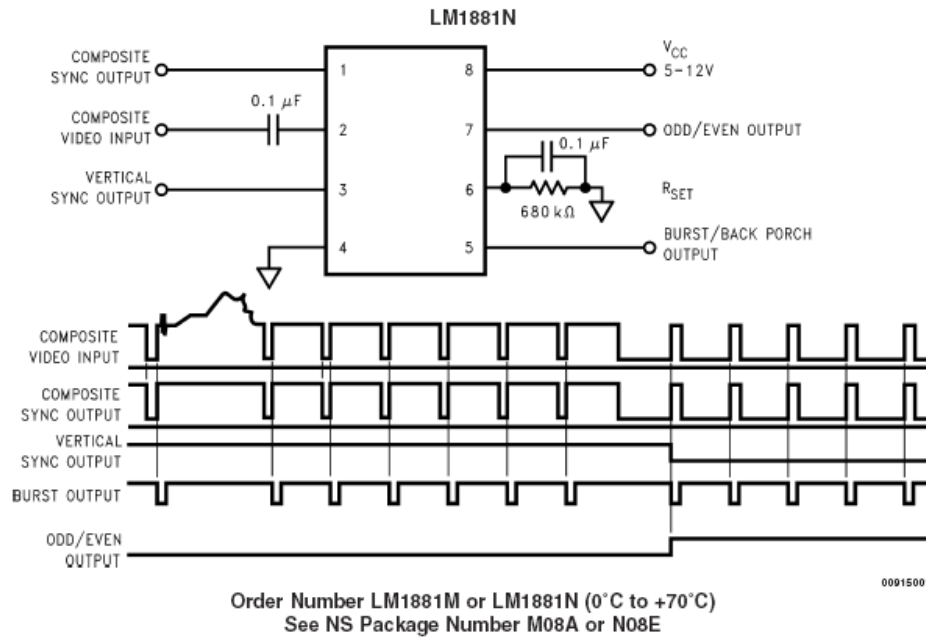


Figura B.12: Diagrama del circuito integrado LM1881N y las señales con las que trabaja.

B.1.5. Integración de los sistemas de visión, iluminación y sincronización

Una vez desarrollado el diseño de cada sistema por separado, resta llevar a cabo la integración y calibración de valores de todos los componentes individuales que fueron anunciados anteriormente.

En la figura B.14 se puede observar el conexionado total, en donde, están incluidos los sistemas de visión, iluminación y sincronización.

B.1 Sistema de visión en el infrarrojo cercano

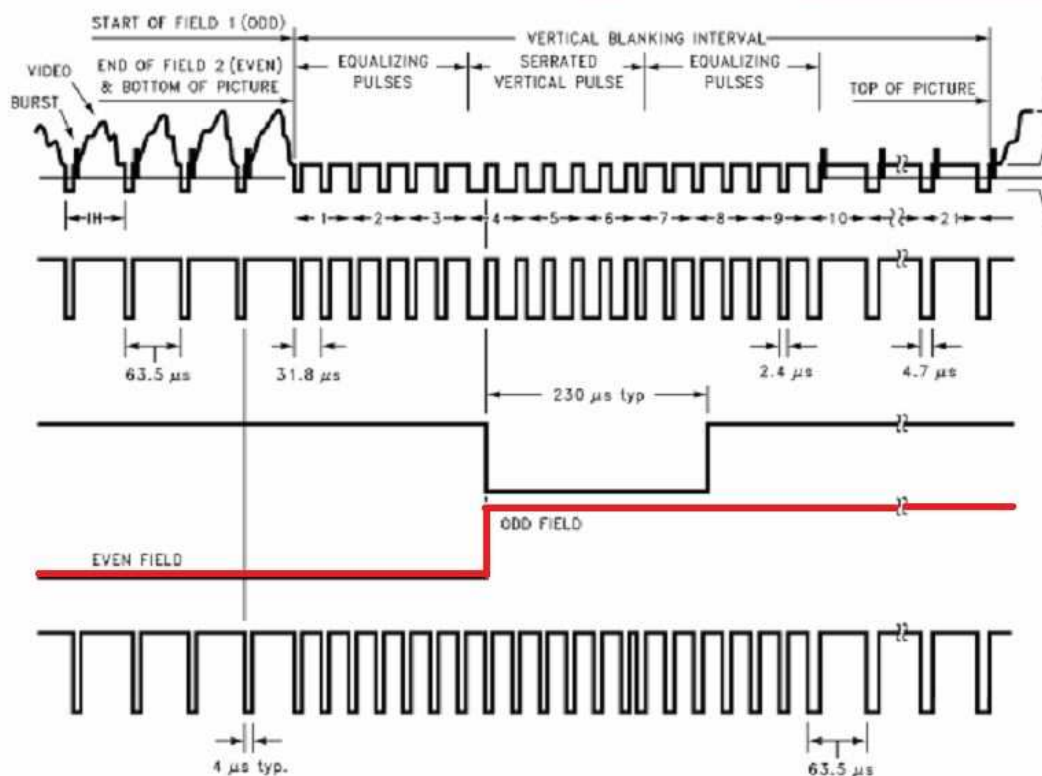


Figura B.13: Señales proporcionadas por el LM1881N. En rojo la señal utilizada para el campo.

B.1.5.1. Integragción del sistema video-sincronización

Dentro del esquema de conexionado, la integración conjunta de los sistemas de visión y sincronización se presenta en el esquema la figura B.15.

Como ya se ha expuesto anteriormente la cámara de video proporciona una señal en video compuesto que se introduce en el sistema de sincronización.

Como se ve en la figura B.16 la señal de video compuesto proporcionada por el sistema de visión tiene un valor máximo de 1 Vpp. Esta será por tanto la señal que se debe decodificar y por tanto usar como entrada del sistema de sincronización.

Esta señal es introducida en el LM1881N que se encargará de obtener la señal de sincronismo del campo par/impar del cuadro de la imagen. Como se ha comentado anteriormente se necesitan dos señales que hagan disparar el encendido del sistema de iluminación alter-

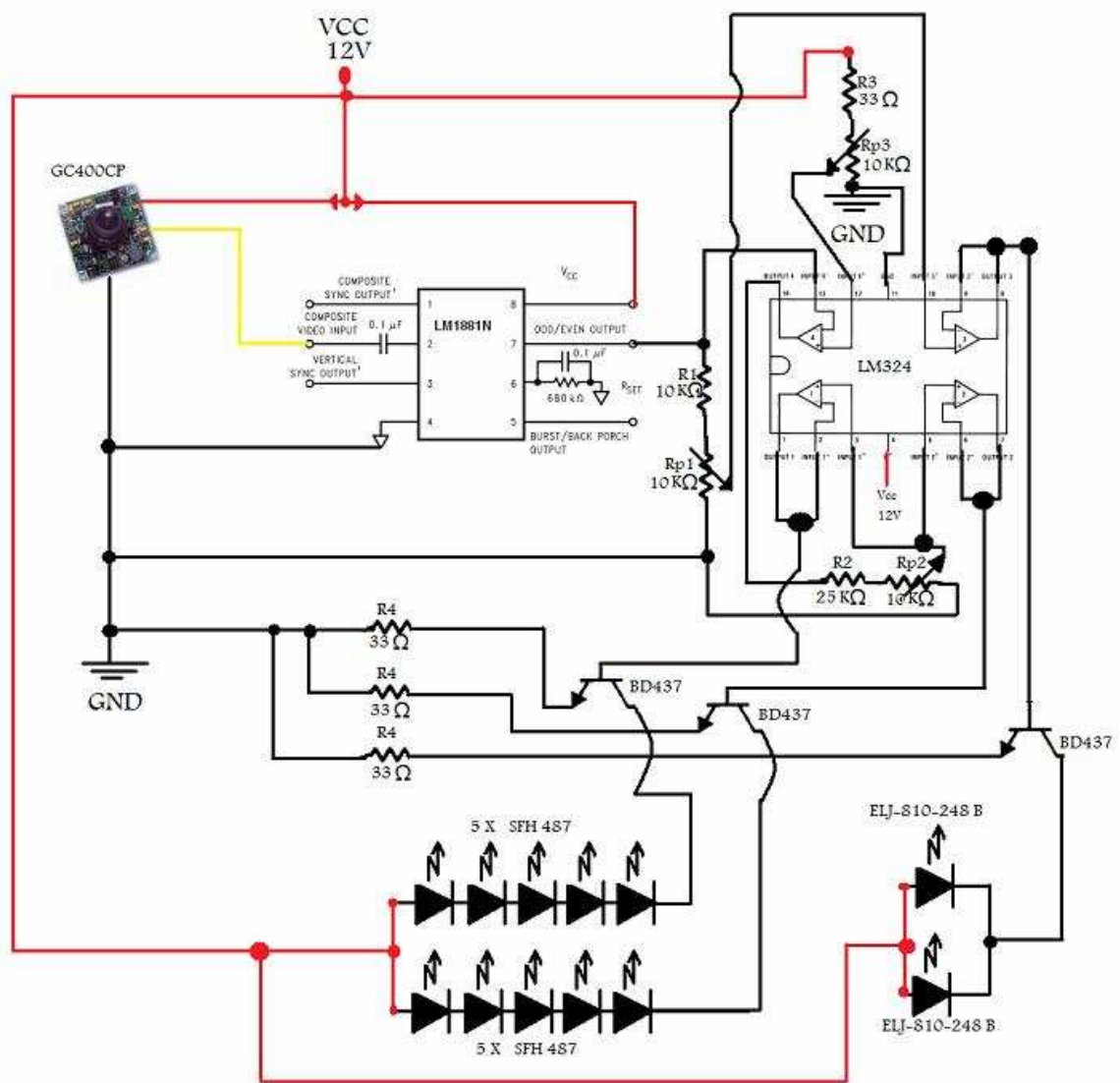


Figura B.14: Esquema del sistema de visión con iluminación infrarroja.

nadamente. La señal que proporciona el LM1881N por la salida de campo par/impar es una señal cuadrada cuya amplitud depende de la tensión de alimentación siendo positiva para el campo impar y cero para el campo par. En la figura B.17 se observa la salida del LM1881N.

Según la hoja de características, para una alimentación de 5 V la salida típica ronda los 4,5 V. Como se observa la configuración de este sistema está alimentado a 12 V. Como tampoco las especificaciones técnicas de estos elemento son precisas lo que hacemos

B.1 Sistema de visión en el infrarrojo cercano

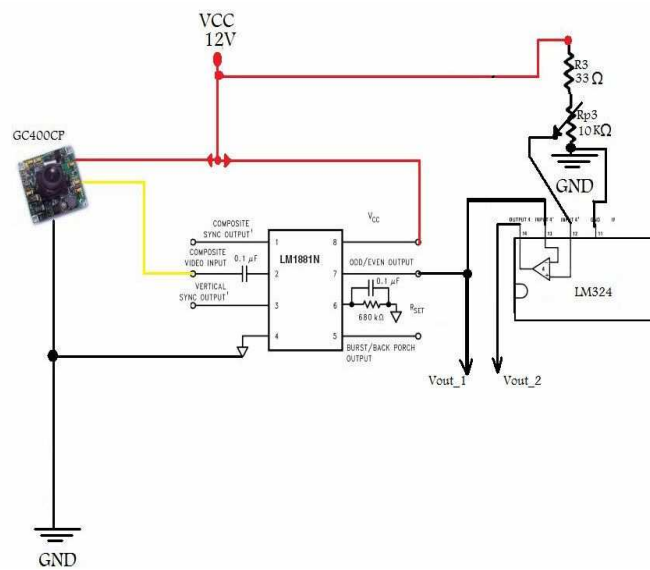


Figura B.15: Conexión de los sistemas de visión y sincronización.



Figura B.16: Señal de video compuesto proporcionado por el sistema de visión.

es obtener la medida de tensión de salida del integrado LM1881N de forma práctica y diseñar los valores y resto de componentes a partir de este punto. Por tanto procedemos a verificar el valor obtenido.

La amplitud obtenida es de 7,5 Vpp y un período de 40 ms que son los 25 Hz esperados. Por tanto se procede a diseñar el resto de componentes desde este punto.

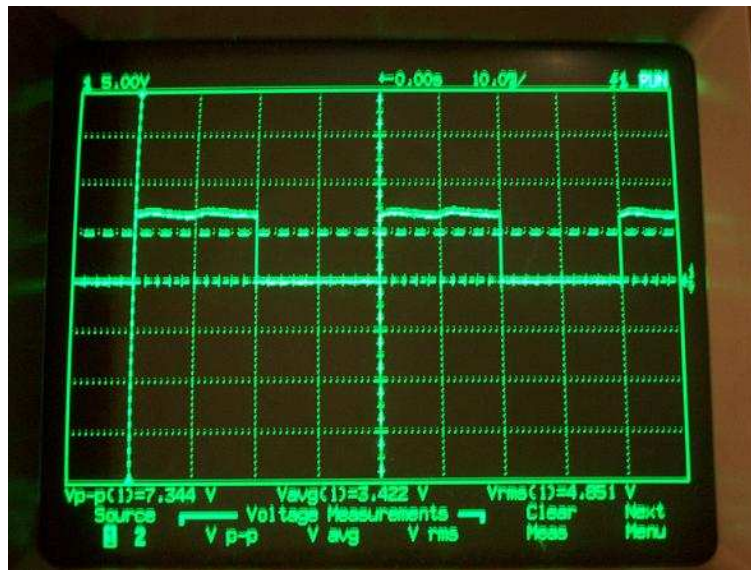


Figura B.17: Señal de sincronismo obtenida en el LM1881N.

Si esta señal controlará el disparo del momento de sincronismo del campo par, se ha de conseguir una señal análoga desfasada 180° para controlar el encendido en el campo impar.

Por este motivo se observa como la señal de campo de Vout_1 se la introduce en el LM324 en configuración inversora para que de esta forma se pueda obtener la misma señal invertida y así conseguir el efecto de alternancia en el encendido de los sistemas de iluminación. El inversor ha sido diseñado con una configuración básica donde la entrada V+ del operacional está controlada por un potenciómetro, de forma que se pueda regular el nivel de disparo que se quiera conseguir. Esta decisión ha sido tomada por si en un futuro se quiere estudiar un nuevo sistema donde no se quiera conseguir alternancia en el encendido, sino una sincronización total en ambas señales (que funcionen a la vez) y así reaprovechar esta cámara. Para el propósito de esta investigación se deja este valor en 5.5 V de forma que a partir de este nivel se dispare. La salida del integrado será la señal de entrada en contratase con una amplitud de 12 V que es el valor de alimentación del LM324.

Se destaca que el tiempo de respuesta del LM324 ante un cambio de amplitud del escalón es de $10 \mu s$ siendo despreciable para el propósito perseguido.

B.1 Sistema de visión en el infrarrojo cercano

Es este punto por tanto donde se tiene dos señales (Vout_1 y Vout_2) con una amplitud de 7,5 V para Vout_1 y 12 V para Vout_2, con período de 25 Hz en contrafase y que serán las que enciendan los sistemas de iluminación de forma alternante.

B.1.5.2. Integreción del sistema de sincronismo-iluminación

Como hemos visto existe un desnivel importante de amplitud entre las dos señales Vout_1 y Vout_2. Estas señales se introducirán en los sistemas de iluminación de forma que su amplitud se pueda regular para de esta forma poder controlar la potencia radiada.

Como tampoco se quiere unos valores de radiación importantes que puedan dañar la estructura del ojo y tampoco distraigan al conductor, se evalua que para una tensión de entrada de cada sistema radiante en 3,5 V se consigue tener estos dos factores cumplidos. No obstante más adelante se expondrá la cantidad de potencia que se radiará.

Por este motivo se ha creado un divisor de tensión para cada señal de forma que sus valores estén comprendidos entre 3,5 V y 0 V.

Se dividirá la integración del sistema de sincronización-iluminación total en dos partes, uno para los dos ELJ-810-248 B y otro para los SFH 487. Empezaremos con la integración de los dos ELJ-810-248 B.

El esquema utilizado para los jumbos viene en la figura B.18.

La señal que controlará el disparo será la que proviene de Vout_1. Recordar que esta señal tenía una amplitud de 7.5 V y un período de 25 Hz. Como se desea poder graduar el nivel de radiación de los dos emisores de luz, lo que se hace es controlar la tensión de entrada del LM324 mediante el uso de un divisor de tensión. El divisor de tensión utilizado es una resistencia de 10 $K\Omega$ y un potenciómetro de 10 $K\Omega$.

Con este divisor s tiene que la amplitud de la señal de entrada estará comprendida entre 3,625 Vpp y 0 V.

Con este rango de amplitud se consigue el siguiente comportamiento del sistema de iluminación:

B.1 Sistema de visión en el infrarrojo cercano

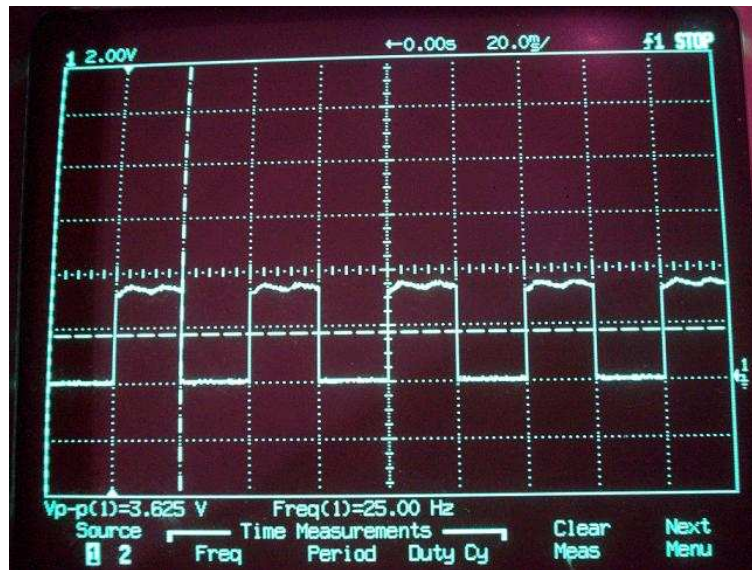


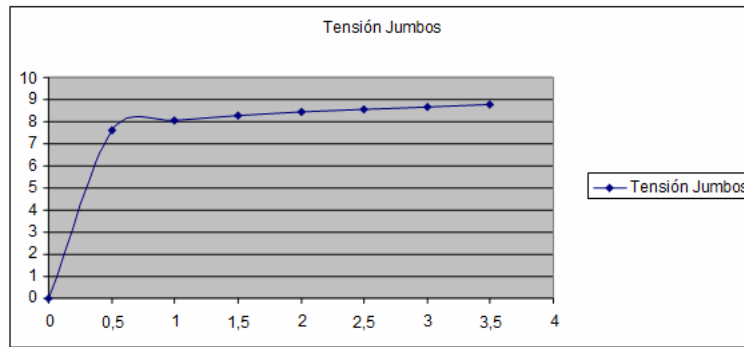
Figura B.19: Señal de control de potencia de los jumbos.

Vout_1(V)	V_jumbos (V)	V_33Ω (V)	I_33(mA)	Vbe(V)
0	0	0	0	0
0,5	7,6	0,069	2,09090	0,5
1	8,08	0,45	13,6363	0,54
1,5	8,29	0,93	28,1818	0,56
2	8,43	1,42	43,03	0,57
2,5	8,56	1,92	58,18	0,58
3	8,67	2,41	73,03	0,6
3,5	8,77	2,89	87,57	0,6

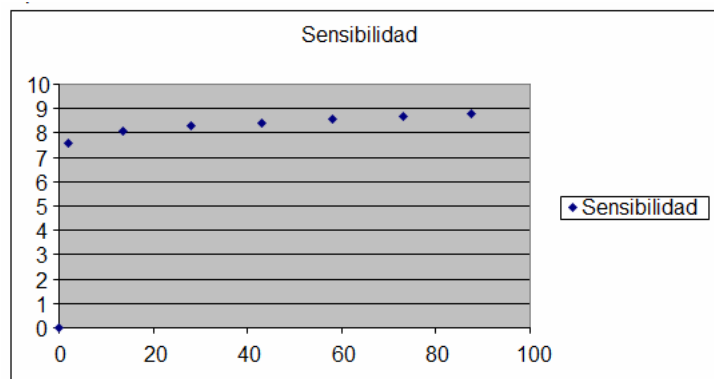
Tabla B.2: Comportamiento del sistema de iluminación para varias tensiones.

Parameter	Test conditions	Symbol	Min	Typ	Max	Unit
Forward voltage	$I_F = 100 \text{ mA}$	V_F		10,5		V
Forward voltage	$I_F = 250 \text{ mA}$	V_F		13		V

Tabla B.3: Voltajes teóricos de los jumbos.



(a)



(b)

Figura B.20: Comportamiento dinámico de los jumbos: (a) tensión, (b) sensibilidad.

V/mA. Si se compara este valor de sensibilidad teórico con el experimental se observa que se obtiene una pendiente de 0,013 V/mA, muy cercano al teórico.

Respecto a la corriente que circula por los ELJ-810-248 y que al fin y al cabo controlarán la potencia de radiación de estos, se puede medirla en la resistencia de potencia de $33\ \Omega$, ya que como se ve se tiene polarizado el transistor en directa, de forma que la corriente de colector y emisor pueden ser consideradas iguales. El resultado obtenido se muestra en la figura B.21.

El comportamiento de la fuente es totalmente lineal, de manera que el potenciómetro controla de forma lineal la cantidad de radiación que los ELJ-810-248 pueden emitir. Se destaca que la corriente que pasa por la resistencia de $33\ \Omega$ está dividida a la mitad en cada

B.1 Sistema de visión en el infrarrojo cercano

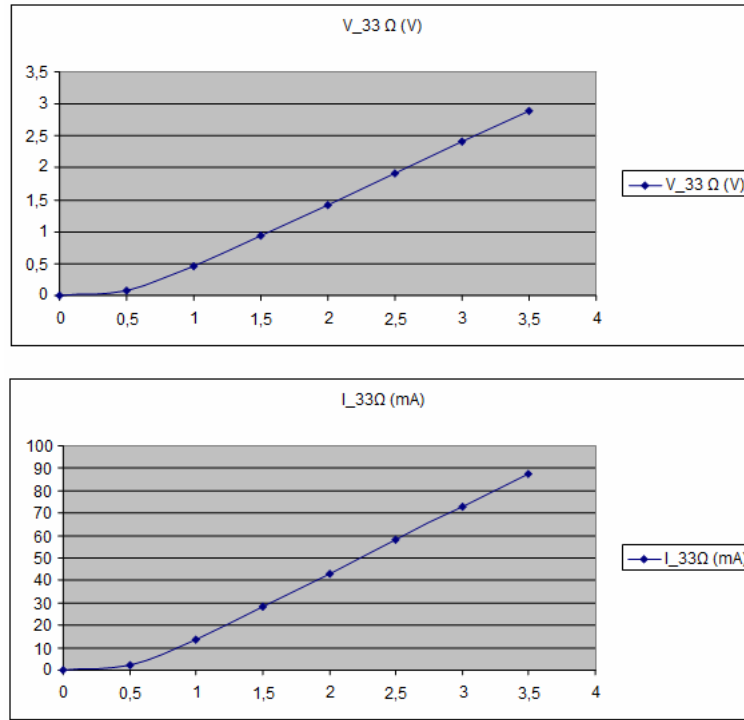


Figura B.21: Comportamiento dinámico fuente de corriente jumbos.

rama de los ELJ-810-248, es decir, por cada ELJ-810-248 va la mitad del valor obtenido en la resistencia de potencia.

Si se entiende un comportamiento lineal en cuanto a la potencia de radiación y utilizando el valor de $I_F = 250$ mA se puede estimar una sensibilidad de 0,875 mW/mA. Con esta sensibilidad y observando un valor máximo de corriente de 45 mA (redondeando a la alta para cada ELJ-810-248), el valor máximo de radiación es de 39,375 mW. Estimando este valor se obtiene a la salida del elemento radiante el valor de 175 mW/cm², sobre una superficie de 0,5 cm x 0,5 cm.

Este valor es justo a la salida del emisor. Este campo radiado se abre con una amplitud de 40 grados, de forma que la potencia se distribuye en el área del spot del receptor.

El siguiente paso será la integración del sistema de sincronización-iluminación correspondiente a cada grupo de 5 leds SFH-487.

Como anteriormente se ha expuesto, el sistema está compuesto de 10 leds SFH 487 que

forman el anillo central que rodea a la lente de la cámara, estos se encuentran divididos en dos grupos compuestos por 5 emisores SFH 487. El diseño de los dos sistemas es idéntico y está compuesto por el esquema de la figura B.22.

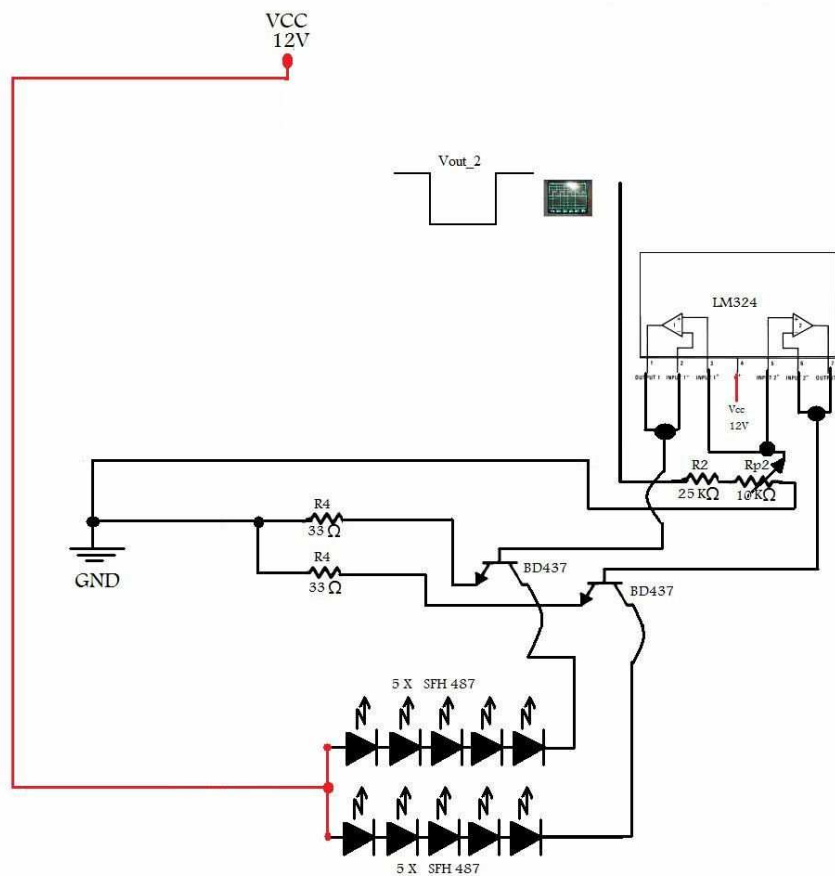


Figura B.22: Conexión del sistema de sincronización-iluminación del grupo compuesto por los emisores que forman el anillo central.

B.1 Sistema de visión en el infrarrojo cercano

Recordar que la señal que controla este sistema es Vout_2, una señal de 12 V y desfasada 180° respecto a Vout_1.

Al igual que en el sistema anterior, la señal será controlada por un potenciómetro de forma que se pueda graduar la intensidad de los elementos radiantes. Para poder disponer del mismo nivel de tensión que en el sistema anterior se crea un divisor resistivo con una primera resistencia de 25 K Ω y un potenciómetro de 10 K Ω de forma que la tensión estará comprendida entre 3,25 V y 0 V.

Por tanto se ha conseguido tener dos señales, Vout_1 y Vout_2, con una amplitud similar y en contraste que controlará el momento de encendido de ambos sistemas radiantes, figura B.23.



Figura B.23: Señales de control de disparo y potencia de los jumbos (arriba) y el anillo central (abajo).

Como se puede observar en el esquema esta señal (Vout_2) controla a ambas fuentes de corriente de forma que siempre se asegura el mismo comportamiento y se consigue que el anillo central se comporte al unísono con este potenciómetro.

El resultado al evaluar el comportamiento dinámico del sistema se presenta en la tabla B.4.

Vout_2(V)	V_leds (V)	V_33Ω (V)	I_33(mA)	Vbe(V)
0	0	0	0	0
0,5	3	0,0984	2,9818	0,5
1	4,8	0,325	9,8484	0,54
1,5	7,5	0,9	27,2727	0,56
2	7,5	1,3	39,39	0,57
2,5	7,5	1,86	56,3636	0,6
3	7,5	2,4	72,7272	0,6
3,5	7,5	2,5	75,75	0,6

Tabla B.4: Comportamiento del sistema de los leds para varias tensiones.

Siendo V_Leds la tensión que cae en un grupo de 5 SFH 487, V_33 Ω la tensión que cae en la resistencia de 33Ω, I_33Ω la corriente que pasa por dicha resistencia y por tanto la que pasa por cada grupo de SFH 487 y por último Vbe el valor de tensión base-emisor del transistor. En la figura B.24 se observa este comportamiento dinámico.

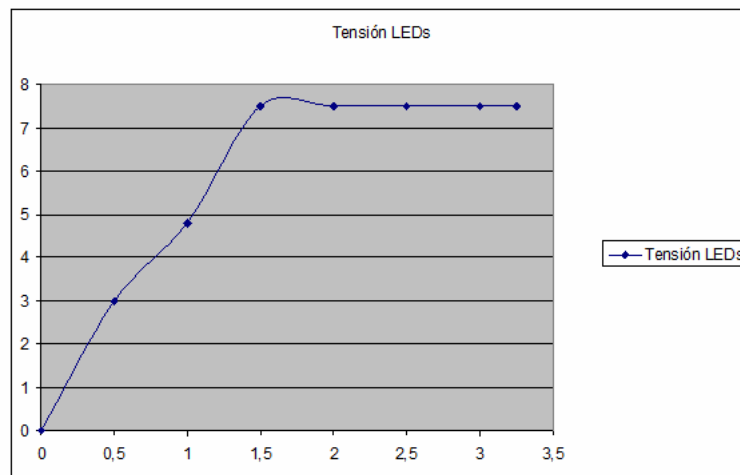


Figura B.24: Comportamiento dinámico del grupo de leds SFH 487.

Se observa como la tensión que cae en el grupo de 5 leds crece dinámicamente hasta estabilizarse en alrededor de los 7,5 V. Como se dijo anteriormente, ésta es la tensión que recae en un grupo de 5 leds, de forma que suponiendo que son ideales y llevan el mismo

B.1 Sistema de visión en el infrarrojo cercano

comportamiento, en cada led recae una tensión de 1,5 V.

B.1.6. Montaje y resultados

Una vez que se tienen diseñados e integrados todos los sistemas se pasa al montaje y pruebas de funcionamiento.

Tanto el módulo de sincronización y las fuentes de corriente que controlan los módulos de iluminación se integran sobre una misma placa que se comunica con la cámara y elementos radiantes mediante un cable.

El soporte de la cámara ha sido diseñado de forma que pueda ser colocado sobre el quitasol del vehículo Ivvi sin sobresalir del tamaño del mismo, evitando pérdidas en el campo de visión del conductor.

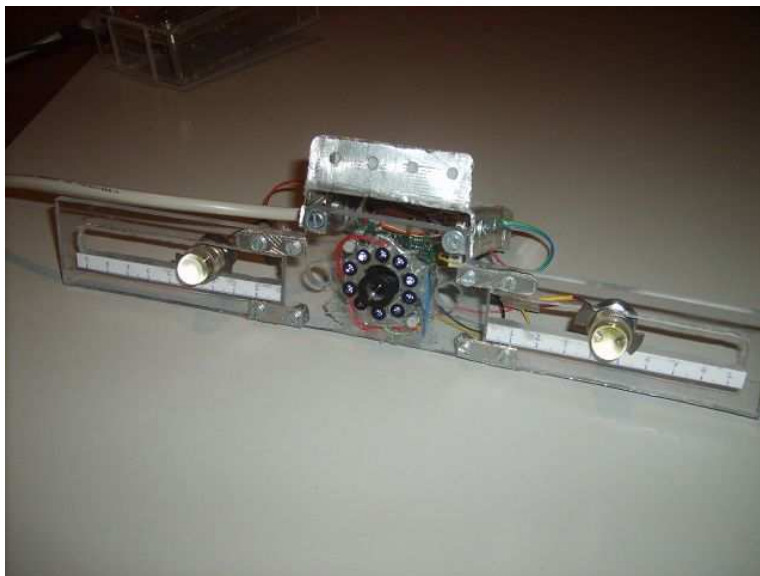


Figura B.25: Sistema de visión con iluminación infrarroja.

Una vez dispuesta la cámara se pasa a evaluar la cantidad de potencia recibida por el conductor y estimar si se sobrepasa lo recomendado que es de 10 mW/cm^2 . Para ello se diseña un sistema que consiste en un diodo PIN BPW34 sensible en la longitud de onda que se está estudiando.

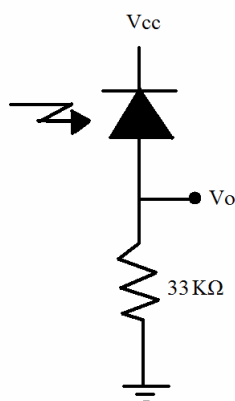


Figura B.26: Esquema de medida de la intensidad radiada.

Este dispositivo ofrece una corriente proporcional a la potencia recibida. Al estimar una longitud de onda de 850 nm la sensibilidad del sistema es de 0,9 W/A.

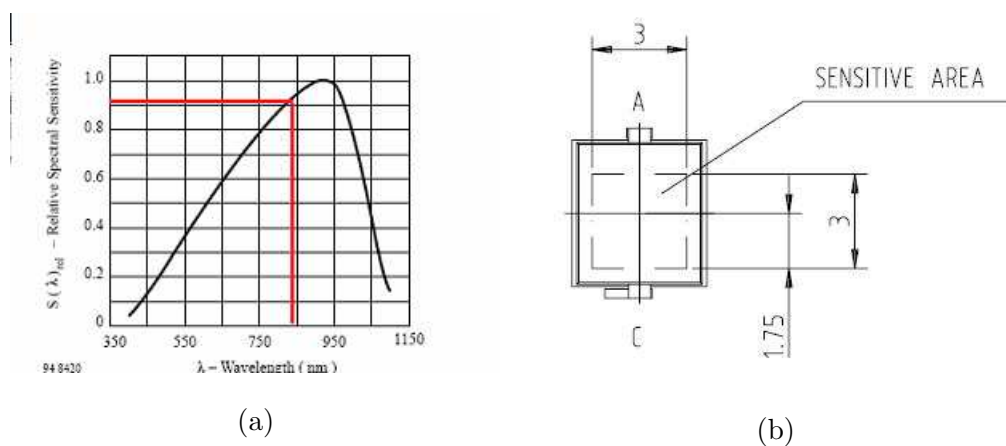


Figura B.27: Sensibilidad BPW34 y área sensible.

Por tanto se va evaluando el valor de corriente obtenido para diferentes distancias. El experimento se realiza desde una distancia mínima de 10 cm hasta una distancia máxima de 70 cm.

Igualmente el área del receptor es un cuadrado de 3 mm de lado, por tanto su superficie es de $0,09 \text{ cm}^2$. Los resultados se presenta en la tabla B.5.

B.1 Sistema de visión en el infrarrojo cercano

Distancia	Vr (V)	Ir (mA)	P (mW)	P/S (mW/cm ²)
10	3,40	0,103030	0,0927	1,0303
20	1,2	0,03636	0,0327	0,3636
30	0,69	0,0209	0,0188	0,209
40	0,48	0,0145	0,01309	0,1454
50	0,28	0,00848	0,007636	0,0848
60	0,2	0,00606	0,005454	0,0606
70	0,1	0,00303	0,002727	0,0303

Tabla B.5: Potencia recibida en función de la distancia.

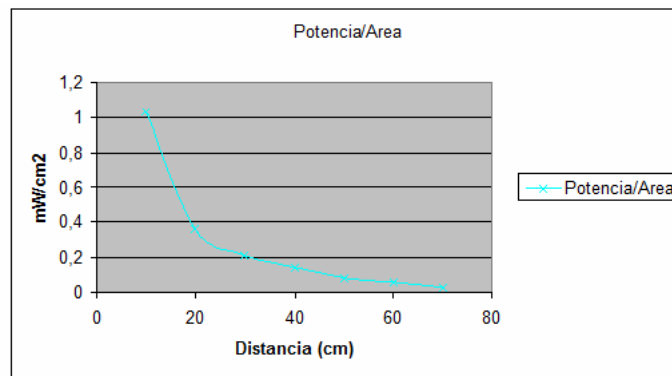


Figura B.28: Potencia recibida por área de superficie en función de la distancia.

Las pérdidas de propagación hacen que se reciba un máximo de $1,03 \text{ mW/cm}^2$, valor que dista mucho de los 10 mW/cm^2 recomendados como umbral máximo de seguridad, por lo tanto, *la seguridad del conductor esta garantizada.*

Conceptos de Estadística y Probabilidad

En esta parte del anexo se explica brevemente y de manera formal la Teoría de Probabilidades, que es una ramificación de las matemáticas que estudian los fenómenos aleatorios. Los términos variables aleatorias, procesos estocásticos y eventos serán definidos para comprender mejor varias de las secciones que componen esta tesis. En este sentido, se incorpora una breve introducción al algoritmo EM (Expectation Maximization) y al algoritmo Condensation.

C.1. Variables aleatorias, espacio muestral, función de distribución acumulada

En general, una variable aleatoria X es una función que va del espacio muestra Ω a los números reales \mathbb{R} donde $X : \Omega \rightarrow \mathbb{R} : w_k \rightarrow x_k$ con $w_k \in \Omega$ y x_k es una realización de la variable aleatoria X .

La función de distribución acumulada (cdf) de una variable aleatoria X , denotada por $F(x) : \mathbb{R} \rightarrow [0, 1]$ es definida por:

$$F(x) = P(X \leq x), \forall x \in \mathbb{R} \tag{C.1}$$

La función $F(x)$ es una cdf si y solamente si cumple las siguientes restricciones: (i) $\lim_{x \rightarrow -\infty} F(x) = 0$ y $\lim_{x \rightarrow \infty} F(x) = 1$; (ii) $F(x)$ es una función no decreciente de x ; (iii)

$F(x)$ es continua a derecha.

C.1.1. Función de densidad de probabilidad

Asociado con una variable aleatoria X y su cdf $F(x)$ está la función de densidad de probabilidad (pdf). La pdf de una variable aleatoria discreta X está dada por

$$p(x) = P(X = x), \forall x \in \mathbb{R} \quad (\text{C.2})$$

y la pdf de una variable aleatoria continua X es la función que satisface

$$F(x) = \int_{-\infty}^x p(t)dt \quad (\text{C.3})$$

para todo $x \in \mathbb{R}$. Para cualquier $a < b$ la probabilidad de que X se encuentre en el intervalo (a, b) es el área bajo la curva de la pdf sobre el intervalo:

$$P(a < x < b) = \int_a^b p(t)dt \quad (\text{C.4})$$

Una función $p(x)$ es una pdf de una variable aleatoria X si y solamente si (i) $p(x) \geq 0$, $\forall x$, (ii) $\int_{-\infty}^{\infty} p(x)dx = 1$.

C.1.2. Valor esperado y varianza

Dos de los parámetros usados para describir o resumir las propiedades de una variable aleatoria X son la media $E(X)$ y la varianza $Var[X]$. La media o valor esperado de X , $\mu = E[X]$ está dado por

$$\mu = E[X] = \sum_k x_k p(x_k) \quad (\text{C.5})$$

en el caso discreto X (provado que $\sum_k |x_k|p(x_k) < \infty$), y

$$\mu = E[X] = \int_{-\infty}^{\infty} xp(x)dx \quad (\text{C.6})$$

en el caso continuo (provado que $\int_{-\infty}^{\infty} |x|p(x)dx < \infty$). La varianza de X está definida por $\sigma^2 = Var[X] = E[(X - \mu)^2]$ que es igual a

$$\sigma^2 = Var[X] = \sum_k (x_k - \mu)^2 p(x_k) \quad (C.7)$$

para el caso discreto, y para el caso continuo

$$\sigma^2 = Var[X] = \int_{-\infty}^{\infty} (x - \mu)^2 p(x) dx \quad (C.8)$$

C.1.3. Distribuciones discretas

Las pdf discretas más comunes son:

- Bernoulli (q):

$$P(X = x|q) = q^x(1 - q)^{1-x}, \quad x = 0, 1, \quad 0 \leq q \leq 1 \quad (C.9)$$

con media $E[X] = q$ y varianza $Var[X] = q(1 - q)$.

- Binomial (n, q):

$$P(X = x|n, q) = C_x^n q^x (1 - q)^{n-x}, \quad x = 0, 1, 2, \dots, n, \quad 0 \leq q \leq 1 \quad (C.10)$$

con la media $E[X] = nq$ y la varianza $Var[X] = nq(1 - q)$.

- Geometrica (q):

$$P(X = x|q) = q(1 - q)^{x-1}, \quad x = 1, 2, \dots, \quad 0 \leq q \leq 1 \quad (C.11)$$

con la media $E[X] = \frac{1}{q}$ y la varianza $Var[X] = \frac{1-q}{q^2}$.

- Poisson(λ):

$$P(X = x|\lambda) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, \dots, \quad 0 \leq \lambda < \infty. \quad (C.12)$$

con la media $E[X] = \lambda$ y la varianza $Var[X] = \lambda$.

C.1.4. Distribuciones continuas

Las pdf continuas más comunes son:

- Uniforme (a, b):

$$p(x|a, b) = \frac{1}{b - a}, \quad a \leq x \leq b \quad (C.13)$$

con media $E[X] = (a + b)/2$ y varianza $Var[X] = (b - a)^2/12$.

- Beta (α, β) :

$$p(x|\alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, 0 \leq x \leq 1 \quad (\text{C.14})$$

con media $E[X] = \alpha/(\alpha + \beta)$ y varianza $Var[X] = \alpha\beta/((\alpha + \beta)^2(\alpha + \beta + 1))$, $\alpha > 0$, $\beta > 0$ y $B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta)$.

- Chi-square (χ_ν^2) :

$$p(x|\nu) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2}, 0 \leq x \leq \infty \quad (\text{C.15})$$

con media $E[X] = \nu$, varianza $Var[X] = 2\nu$ y ν son los grados de libertad.

- Norma (μ, σ^2) :

$$p(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)} \quad (\text{C.16})$$

con media $E[X] = \mu$ y varianza $Var[X] = \sigma^2$.

C.1.5. Densidad normal multivariante

Sea X un vector de variables aleatorias con media $\mu_X \in \mathbb{R}^n$ y matriz de covarianza $\Sigma_{XX} = E[(X - \mu_X)(X - \mu_X)^T]$, entonces

$$\begin{aligned} E[AX + B] &= A\mu_X + B \\ Var[AX + B] &= A\Sigma_{XX}A^T \end{aligned} \quad (\text{C.17})$$

La función de densidad normal multivariante viene dada por

$$p(X|\mu_X, \Sigma_{XX}) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma_{XX})}} \exp\left(-\frac{1}{2}(x - \mu_X)^T \Sigma_{XX}^{-1} (x - \mu_X)\right) \quad (\text{C.18})$$

que se denota como $X \sim N(\mu_X, \Sigma_{XX})$.

C.2. Introducción al Algoritmo EM

El algorithmo EM (Expectation Maximization) propuesto por Dempster *et al.* en 1.977 [2] es utilizado para estimar de manera iterativa los parámetros de las funciones de distribución de alta complejidad que tienen la característica de presentar datos incompletos u ocultos.

C.2.1. Descripción del algoritmo EM

El algoritmo EM puede aplicarse en muchas situaciones en las que se desea estimar un conjunto de parámetros θ que describen una distribución de probabilidad subyacente, dada únicamente una parte observada de los datos completos producidos por la distribución. En general, supongamos que en cada realización del experimento aleatorio se observa un parámetro z_i y existe un parámetro oculto x_i . Denotando entonces por $Z = \{z_1, \dots, z_m\}$ al conjunto de datos observados en m realizaciones del experimento, por $X = \{x_1, \dots, x_m\}$ al conjunto de datos no observados y por $Y = Z \cup X$ al conjunto completo de datos. Los datos X pueden ser considerados como variables aleatorias cuya distribución de probabilidad depende de los parámetros θ y los datos observados Z . De la misma forma, Y es una variable aleatoria que está definida en términos de la variable aleatoria X . Sea h la hipótesis actual de los valores de los parámetros θ , y sea h_1 la hipótesis revisada que se estima en cada iteración del algoritmo EM.

El algoritmo EM busca la hipótesis h_1 que maximiza la esperanza $E[\ln(p(Y|h_1))]$, siendo $p(Y|\theta)$ la distribución de probabilidad que define Y y que depende de los parámetros desconocidos θ . Esta distribución de probabilidad define la verosimilitud de los datos completos Y dada una hipótesis h_1 de los parámetros ocultos. Al maximizar el logaritmo de la distribución se maximiza la verosimilitud. Se introduce el valor esperado $E[\ln(p(Y|h_1))]$ debido a que el conjunto completo de datos Y es una variable aleatoria. Dado que el conjunto completo de datos Y contiene datos X no observados, se deben considerar todos los posibles valores de X , ponderándolos según su probabilidad. En otras palabras se calcula el valor esperado $E[\ln(p(Y|h_1))]$ sobre la distribución de probabilidad que gobierna la variable aleatoria Y . Esta distribución está determinada por los valores observados Z más la distribución de los valores no observados X .

En general, la distribución de Y es desconocida porque está determinada por los parámetros θ que se intenta estimar. Por ello, el algoritmo EM usa la hipótesis actual h para estimar la distribución de Y . Se define entonces una función $Q(h|h_1)$ que proporciona $E[\ln(p(Y|h_1))]$ como una función de h_1 , bajo la suposición de que $\theta = h$ y dado el

conjunto de observaciones Z del conjunto completo de datos Y

$$Q(h_1|h) = E[\ln(p(Y|h_1))|h, Z] \quad (\text{C.19})$$

en la función $Q(h_1|h)$ se supone que la hipótesis h y los datos observados Z tienen unos valores fijos y que éstos definen la distribución de probabilidad de las variables ocultas X . La distribución de probabilidad de Y definida por Z y h es la que se utiliza para calcular $E[\ln(p(Y|h_1))]$ para una hipótesis cualquiera h_1 .

En su forma general, el algoritmo EM repite los siguientes pasos hasta converger.

1. Paso de estimación (E): Calcular $Q(h_1|h)$ utilizando la hipótesis actual h y los datos observados Z para estimar la distribución de probabilidad Y

$$Q(h_1|h) \leftarrow E[\ln(p(Y|h_1))|h, Z] \quad (\text{C.20})$$

2. Paso de maximización (M): Sustituir h por la hipótesis h_1 que maximiza la función Q

$$h \leftarrow \arg \max_{h_1} Q(h_1|h) \quad (\text{C.21})$$

C.2.2. Estimación del modelo paramétrico del color de la piel

Esta aplicación proviene del trabajo de Flores *et. al.* [92] donde se utiliza el algoritmo EM para estimar los parámetros de la función de distribución de probabilidad del color de piel en el espacio dicromático HS . El modelo de color de piel utiliza la función mezcla de gaussianas y su función de distribución de probabilidad dado el color de piel viene dado por la ecuación (C.22):

$$p(x|\theta, piel) = \sum_{k=1}^m w_k \phi_2(x, \theta_k) \quad (\text{C.22})$$

donde m es el número de componentes en el modelo, $\theta = (\theta_1, \dots, \theta_m)$ es el vector de parámetros (desconocidos), w_k es el peso de la componente k con la restricción $\sum_{k=1}^m w_k = 1$, $\phi_2(\cdot)$ es la función de distribución de probabilidad normal bivalente.

Para este ejemplo, $m = 2$, $x \in R^2$ y θ_k es el vector de medias y la matriz de varianza-covarianza de la componente k y, finalmente, $k = 2$. Sobre una muestra aleatoria del color

C.2 Introducción al Algoritmo EM

de piel se ajusta el modelo antes propuesto, los resultados numéricos se presentan en la tabla C.1 mientras que la figura C.1 presenta el gráfico de la función de distribución del modelo de color de piel y en la figura C.2 se aprecia el filtraje de una imagen con este modelo.

Modelo Mezcla de Gaussianas		
k	Media	Varianza - Covarianza
1	[105.92, 120.49]	[20.8, -40.6, 1795.6]
2	[98.93, 79.65]	[117.89, 33.84, 971.07]

Tabla C.1: Resultados numéricos del algoritmo EM al ajustar el modelo del color de piel sobre el espacio dicromático HS.

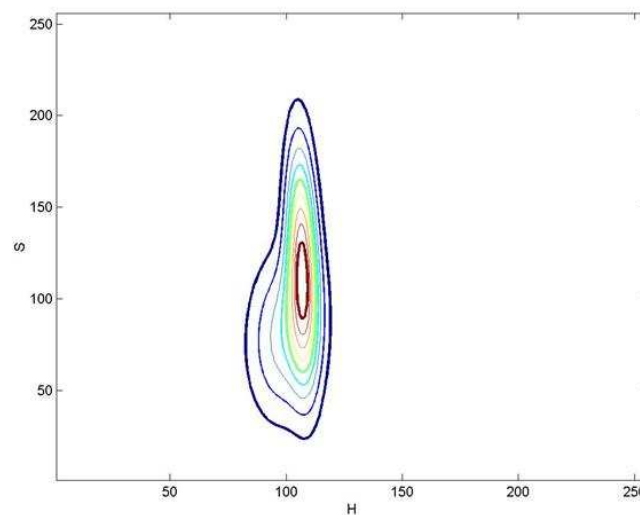
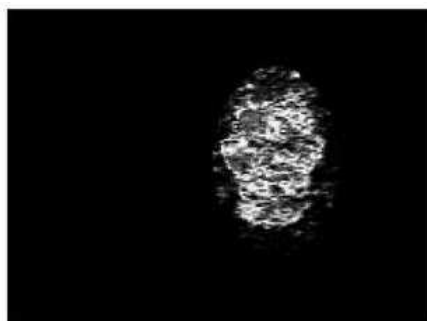


Figura C.1: Gráfico de contorno del modelo del color de piel obtenido con el algoritmo EM.



(a)



(b)

Figura C.2: Resultados del modelo de color de piel, (a) Imagen en color, (b) imagen de probabilidad del color de piel.

Bibliografía

- [1] Branzan A., Widsten B., Wang T., Lan J., and Mah J. A computer vision-based system for real-time detection of sleep onset in fatigued drivers. *IEEE, Intelligent Vehicles Symposium*, pages 25–30, June 2008.
- [2] Dempster A., Laird A., and Rubin D. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 1977. Series B 39, 1 38.
- [3] Doucet A., Freitas N. de, and Gordon N. *Sequential Monte Carlo Methods in Practice*, volume 1. Springer-Verlag, 2001.
- [4] Heitmann A., Guttkuhn R., Aguirre A., Trutschel U., and Moore-Ede M. Technologies for the monitoring and prevention of driver fatigue. *Proceedings of the First International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design*, pages 81–86.
- [5] Isard M. A. *Visual Motion Analysis by Probabilistic Propagation of Conditional Density*. PhD thesis, Department of Engineering Science, University of Oxford, September 1998.
- [6] Turk M. A. and Pentland A. Face recognition using eigenfaces. *IEEE The Media Laboratory (MIT)*, 2001.
- [7] Vicente Ripoli M. A. *Reconocimiento y Búsqueda Selectiva de Objetos en Entornos no Estructurados Mediante Apariencia Multicanal y Multiescala*. PhD thesis, Inge-

niería Electrónica, Departamento de Ingeniería de Sistemas Industriales, Universidad Miguel Hernandez, 2006.

- [8] Vuckovic A., Popovic D., and Radivojevic V. Artificial neural network for detecting drowsiness from EEG recordings. *Neural Network Applications in Electrical Engineering*, pages 155–158, September 2002.
- [9] Gökberk B., Akarun L., and Alpaydin E. Gabor wavelet based pose estimation for face recognition. 2001. <http://www.cmpe.boun.edu.tr/~gokberk/iscis2001.pdf>.
- [10] Ristic B., Arulampalam S., and Gordon N. *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, volume 1. Artech House, 2004.
- [11] Chang C. and Lin C. *LIBSVM: a library for support vector machines*, 2001. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [12] Chang B. C., Lim J. E., Kim H. J., and Seo B. H. A study of classification of the level of sleepiness for the drowsy driving prevention. *IEEE, SICE Annual Conference*, pages 3084–3089, September 2007.
- [13] Feng G. C. Yuen P. C. Multi-cues eye detection on gray intensity image. *Pattern Recognition*, 1(34):1033–1046, January 2001.
- [14] Hilario C. *Detección de Peatones en el Espectro Visible e Infrarrojo para un Sistema Avanzado de Asistencia a la Conducción*. PhD thesis, Departamento de Ingeniería de Sistemas y Automática, Universidad Carlos III de Madrid, Octubre 2008.
- [15] Hilario C., Collado J. M., Armingol J., and Escalera A. de la. Pedestrian detection for intelligent vehicles based on active contour models and stereo vision. *Lecture Notes in Computer Science*, 3643:537–542, 2005.
- [16] Hilario C., Collado J. M., Armingol J., and Escalera A. de la. Pedestrian detection for intelligent vehicles based on active contour models and stereo vision. *10th. International Workshop on Computer Aided Systems Theory*, February 2005. Gran Canaria, Spain.

- [17] Hilario C., Collado J. M., Armingol J., and Escalera A. de la. Visual perception and tracking of vehicles for driver assistance systems. *IEEE Intelligent Vehicle Symposium*, pages 94–99, June 2006.
- [18] Jiangwei C., Lisheng J., Lie G., Keyou G., and Rongben W. Driver’s eye state detecting method design based on eye geometry feature. *IEEE, Intelligent Vehicles Symposium*, pages 357–362, June 2004.
- [19] Jiangwei C., Lisheng J., Lie G., Keyou G., and Rongben W. A monitoring method of driver mouth behaviour based on machine vision. *IEEE, Intelligent Vehicles Symposium*, pages 351–356, June 2004.
- [20] Liu C. Gabor-based kernel pca with fractional power polynomial models for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligent*, 26(5):572–582, May 2004.
- [21] Needham J. C. *Tracking and Modelling of Team Game Interactions*. PhD thesis, School of Computing, The University of Leeds, October 2003.
- [22] Rasmussen C. and Hager G. D. Probabilistic data association methods for tracking complex visual objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):560–576, June 2001.
- [23] Volvo Cars. Control de alerta para el conductor. Website, 2008. <http://www.volvocars.com>.
- [24] Portal Ciencia. Accidentes de tráfico: Una lacra de la sociedad moderna. Website, 2007. <http://www.portalciencia.net/vdc/trafico.htm>.
- [25] Sólo Ciencia. 23 de mayo, el día que la poblacion urbana se volvio más urbana que rural. Website, 2007. <http://www.solociencia.com/antropologia/07062806.htm>.
- [26] El Comercio. Sistema de monitorización del nivel de alerta humano basado en ondas cerebrales. Prensa escrita, Octubre 1, 2006. Cuaderno 2: Tecnología.

-
- [27] Grest D. and Koch R. Real time multi-camera person tracking for immersive environments. *IEEE, 6th. Workshop on Multimedia Signal Processing*, pages 387–390, October 2004.
- [28] Escalera A. de la. *Visión por Computador, Fundamentos y Métodos*, volume 1. Prentice Hall, Pearson Educación, Madrid, 2001.
- [29] Escalera A. de la, Salichs M. A., Moreno L., and Puente E. A. Neural traffic signal recognition for autonomous vehicles. *IEEE International Conference on Industrial Electronics, Control, Instrumentation and Automation*, September 1994.
- [30] Escalera A. de la, Moreno L., Salichs M. A., and Armingol J. M. Traffic sign perception in road environments. *International Symposium on Automotive Technology and Automation*, November 1994.
- [31] Escalera A. de la, Moreno L., and Armingol J. M. Road traffic sign detection and classification. *IEEE Transactions on Industrial Electronics*, 44(6):848–859, June 1997.
- [32] Escalera A. de la, Armingol J. M., and Salichs M. A. Traffic sign detection for driver support systems. *3rd. Conference on Field and Service Robotics, Espoo*, June 2001.
- [33] Escalera A. de la, Armingol J. M., Rodriguez F., and Pastor M. Visual sign information, extraction and identification by deformable models for intelligent vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 5(2):57–68, 2004.
- [34] Escalera A. de la, Armingol J. M., and Pastor M. Traffic sign recognition and analysis for intelligent vehicles. *Image and Vision Computing*, 11(3):247–258, 2003.
- [35] Organizacion Panamericana de la Salud. Por una mejor seguridad vial en las américas. Website, 2007. <http://www.paho.org/English/DD/PIN/seguridadvial.pdf>.
- [36] Accidentes de tráfico. Campaña de sensibilización. Website, 2007. <http://pantha-rei.blogspot.com/>.
- [37] Dirección General de Tráfico. Salud vial. Website, 2004. www.dgt.es/revista/num164/pdf/num164-2004-Salud.pdf.

- [38] Dirección General de Tráfico. Estudio ARAG 2.008: La velocidad en el tráfico. Website, 2008. www.dgt.es/was6/portal/contenidos/documentos/seguridad_vial/estudios_informes/Estudio_Velocidad_ARAG.pdf.
- [39] Laboratorio del sueño. Diagnóstico y tratamiento a pacientes con problemas para dormir. Website, 2007. <http://www.somno.org/laboratorio/proceso.php?estilo=2&accion=info&pag=1&id=10>.
- [40] Ayoob E., Steinfeld A., and Grace R. Identification of an appropriate drowsy driver detection interface for commercial vehicle operations. *Proceedings of 47th Annual Meeting the Human Factors and Ergonomics Society*, October 2003.
- [41] Boser B. E., Guyon I., and Vapnik V. A training algorithm for optimal margin classifiers. In *Fifth Annual Workshop on Computational Learning Theory, Proceedings of*, pages 144–152, 1992.
- [42] Koller-Meier E. and Ade F. Tracking multiple objects using the condensation algorithm.
- [43] Murphy-Chutorian E. and Trivedi M. Head pose estimation in computer vision: A survey. *IEEE : Transactions on Pattern Analysis and Machine Intelligence*, 31(4):607–626, April 2009.
- [44] EcoDiario. Unas 1500 personas mueren al año en accidentes de tráfico relacionados al alcohol. Website, 2008. <http://ecodiario.eleconomista.es/gente-estilo/noticias/659798/08/>.
- [45] Red en medicina del sueño. Trastornos del sueño. Website, 2007. <http://www.rems.com.ar/somnoliencia.htm>.
- [46] Real Academia Española. Diccionario de la lengua española, vigésima segunda edición. Website, 2008. <http://www.rae.es>.
- [47] Bloemkolk F., Lijster J. de, and Gelderen M. van. ITS strategy: the japanese formula for success. Study to promote ITS implementation in the Netherlands. Technical

-
- report, International Affaris Office, Ministry of Transportation, Public Works and Water Management, July 2007.
- [48] Castrillón Santana M. F. *Sobre la Detección en Tiempo Real de Caras en Secuencias de Video. Una Aproximación Oportunista*. PhD thesis, Departamento de Informática y Sistemas, Universidad de las Palmas de Gran Canaria, Diciembre 2002.
- [49] Cootes T. F., Taylor C. J., Cooper D H., and Graham J. Active shape models-their training and application. *Computer Vision and Imange Understanding*, 61(1):38–59, January 1995.
- [50] Küçükay F. and Bergholz J. Driver assistant systems. *Lectures of Institute of Automatic Engineering*, 2005.
- [51] Driving Related Facts and Figures. Driving related facts and figures. Website, 2007. <http://www.driveandsurvive.co.uk/cont5.htm>.
- [52] National Sleep Foundation. Drowsy driving org. Website, 2007. <http://www.sleepfoundation.org/site/c.huIXKjM0IxF/b.2418857/k.A5A7/DrowsyDrivingorg.htm>.
- [53] Bradski G. Computer vision face tracking for use in a perceptual user interface. *Intel Technical Journal*, 2:1–15, 1998.
- [54] Daugman J. G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional cortial filters. *J. Optical Soc. Am.*, 2(7):1160–1169, 1985.
- [55] García G., Zikos G., and Tziritas G. Wavelet packet analysis for face recognition. *Elsevier Image and Vision Computing*, 18:289–297, 2000.
- [56] Genton M. G. *Skew-Elliptical Distributions and their Applications*. Chapman & Hall/Crc Press, 2004.
- [57] Kitagawa G. Monte carlo filter and smoother for non-gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, 1(5):1–25, 1996.

- [58] Longhurst G. Understanding driver visual behaviour. *Seeing Machine Pty Limited*.
- [59] Bosh Group. Car multimedia. Website, 2007. <http://www.bosch.com/content/language2/html/index.htm>.
- [60] Gu H., Ji Q., and Zhu Z. Active facial tracking for fatigue detection. *IEEE, Proceedings of the Applications of Computer Vision (WACV'02)*, (1):137–142, June 2002.
- [61] Kashima H., Hongo H., Kato K., and Yamamoto K. A robots iris detection method of facial and eye movement.
- [62] Liu H., Wu Y., and Zha H. Eye states detection from color facial image sequence. *Proceedings of SPIE 4875*, 693, 2002.
- [63] Yang M. H., Kriegman J., and Ahuja N. Detecting faces in images: A survey. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 1(24):34–58, January 2002.
- [64] Guyon I., Gunn S., Nikravesh M., and Zadeh L. A. *Feature Extraction Foundations and Applications*, volume 207 of *Studies in Fuzziness and Soft Computing*. Springer-Verlag Berlin Heidelberg, 2006.
- [65] Kim I., Lee S., Ahn S., Kwon Y., and Kim H. 3d tracking of multi-objects using color and stereo for hci. *IEEE, International Conferences on Image Processing*, 3:278–281, 2001.
- [66] Pitas I. *Digital Image Processing Algorithms and Applications*. A Wiley-Interscience Publication. John Wiley & Sons, Inc., 2000.
- [67] Information Society Technologies (ADASE II). Advanced Driver Assistance Systems in Europe. Technical report, 2007. <http://www.adase2.net>.
- [68] Attention Technologies Inc. Driver fatigue is the number one cause of catastrophic truck accidents. Website, 2005. <http://www.attentiontechnology.com/index.html>.
- [69] IST. Adaptive integrated driver-vehicle interface. Final Report AIDE-D4.2.3a.doc, Information Society Technology (IST) Programme, March 2004.

-
- [70] Armingol J., Aubert D., Azouaoui O., Bauk S., Boussard C., Carrasco J., Chapuis R., Cheng M., Collado J., Escalera A. de la, Douret J., Flores M., Han C., Hauiere N., Hilario C., Huang H., Kodogiannis V., Kulic R., Labayrade R., Lam S., Li Y., Lygouras J., So A., Ieng S., Sridharan K., Tarel J., Wang C., Xu Q., Yazid H., and Yu S. *Autonomous Robots Research Advances*. Nova, 2008.
- [71] Collado J., Hilario C., Escalera A. de la, and Armingol J. Adaptative road line detection and classification. *Lecture Notes in Computer Science*, 4179.
- [72] Collado J., Hilario C., Escalera A. de la, and Armingol J. Model based vehicle detection for intelligent vehicles. *IEEE Intelligent Vehicle Symposium*, pages 572–577, June 2004.
- [73] Collado J., Hilario C., Escalera A. de la, and Armingol J. Detection and classification of road lanes with a frequency analysis. *IEEE Intelligent Vehicle Symposium*, pages 77–82, June 2005.
- [74] Collado J., Hilario C., Armingol J. M., and Escalera A. de la. On board camera perception and tracking of vehicles. *1st. International Workshop on Robot Vision*, pages 57–66, March 2007.
- [75] Satake J. and Shakunaga T. Multiple target tracking by appearance-based condensation tracker using structure information. *IEEE, Proceedings of the 17th International Conference on Pattern Recognition (ICPR04)*, 3:294–297, August 2004.
- [76] Suykens J., Van Gestel T., Brabanter J. de, Moor B. de, and Vandewalle J. *Least Square Support Vector Machines*. K. U. Leuven, 2002.
- [77] Mattsson K. In-vehicle prediction of truck driver sleepiness, lane position variables. Master’s thesis, Department of Computer Science and Electrical Engineering, Luleå University of Technology, 2007.
- [78] Bergasa L., Nuevo J., Sotelo M., and Vázquez M. Real-time system for monitoring driver vigilance. *IEEE, Intelligent Vehicles Symposium*, 1(2), Junio 14-17 2004.

- [79] Bergasa L., Nuevo J., Sotelo M., Barea R., and Lopez E. Real-time system for monitoring driver vigilance. *IEEE, Transactions on Intelligent Transportation Systems*, 7(1):63–77, March 2006.
- [80] Fletcher L., Petersson L., and Zelinsky A. Driver assistance systems based on vision in and out of vehicles. *IEEE, Proceedings of Intelligent Vehicle Symposium*, pages 322–327, 2003.
- [81] Fletcher L., Petersson L., Bames N., Austin D., and Zelinsky A. A sign reading driver assistance system using eye gaze. *IEEE International Conference on Robotics and Automation*, 2005.
- [82] Hagenmeyer L. *Development of a Multimodal, Universal Human-Machine-Interface for Hypovigilance-Management-Systems*. PhD thesis, Mechanical Engineering, University of Stuttgart, Institute for Human Factors and Technology Management, August 2007.
- [83] Huang L. L., Shimizu A., and Kobatake H. Classification-based face detection using gabor filter features. *IEEE, Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04)*, July 2004.
- [84] Martinez W. L. and Martinez A. R. *Computational Statistics Handbook with Matlab*. Chapman & Hall/CRC, 2002.
- [85] Phung S. L., Chai D., and Bouzerdoum A. A universal and robust human skin color model using neural networks. *IEEE International Joint Conference on Neural Networks*, 4:2844–2849, 2001.
- [86] Plung S. L., Chai D., and Bouzerdoum A. Skin colour based face detection. *Seventh Australian and New Zealand Intelligent Information System Conference*, 1(1):171–176, November 2001.
- [87] Vlacic L., Parent M., and Harashima F. *Intelligent Vehicle Technologies*. A division of Reed Educational and Professional Publishing Ltda. Library of Congress Cataloguing in Publication Data, 2001.

-
- [88] Hanmi Industries Ltd. Drowsy truck drivers. Website, 2005. <http://www.gohanmi.com/NREC-COPILOT.htm>.
- [89] Lukor. *Noticias y Motor*, 2007-2008. <http://www.lukor.com/viajes/noticias/04071228.html>.
- [90] Armingol J. M., Escalera A. de la, Hilario C., Collado M., Carrasco J., Flores M., Pastor J., and Rodriguez J. Ivvi: Intelligent vehicle based on visual information. *Robotics and Autonomous Systems*, 55(12):904–916, 2007.
- [91] Bergasa Pascual L. M. *Seguimiento Facial Mediante Visión Artificial Orientado a la Ayuda a la Movilidad*. PhD thesis, Escuela Politécnica, Departamento de Electrónica, Universidad de Alcalá, 1999.
- [92] Flores M., Armingol J., and Escalera A. de la. New probability models for face detection and tracking in color images. *IEEE International Symposium on Intelligent Signal Processing*, pages 871–876, October 2007.
- [93] Isard M. and Blake A. Condensation - conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1):5–28, 1998.
- [94] Isard M. and Blake A. Condensation - conditional density propagation for visual tracking. *Intel Journal of Computer Vision*, 1998. University of Oxford.
- [95] Kaneda M., Obara H., and Nasu T. Adaptability to ambient light changes for drowsy driving detection using image processing. (98-S2-P-15):486–491.
- [96] Kutila M. *Methods for Machine Vision Based Driver Monitoring Applications*. PhD thesis, Tietotalo Building, Auditorium TB104, December 2006.
- [97] Lu M., Wevers K., van der Heijden R., and Heijer T. ADAS applications for improving traffic safety. *IEEE International Conference on Systems, Man and Cybernetics*, 4:3995–4002, October 2004.
- [98] Ray M., Kumar Subudhi P., Tiwari U., Vinod N., and Chandel R. Face recognition using gabor wavelets. *IEEE, Conference on Signals, Systems and Computers ACSSC'06*, pages 593–597, November 2006.

BIBLIOGRAFÍA

- [99] Rizon M., Firdaus M., Saad P., Yaacob S., and Mamat M. Face recognition using eigenfaces and neural networks. *American Journal of Applied Sciences*, 2:1872–1875, June 2006.
- [100] Rizon M. and Kawaguchi T. Automatic eye detection using intensity and edge information. *IEEE, Proceedings TENCON 2000*, 2:415–420, September 2000.
- [101] Zhou M. and Wei H. Face verification using gabor wavelets and adaboost. *IEEE, 18th. International Conference on Pattern Recognition ICPR06*, 1:404–407, 2006.
- [102] Seeing Machines. Visionary technology. Website, 2007. <http://www.seeingmachines.com/index.htm>.
- [103] Super Motor. Accidentes de tráfico: ser responsable, la mejor prevención. Website, 2004. <http://www.supermotor.com/revista/seguridad-vial/249993/accidentes-trafico:-ser-responsable-mejor-prevencion.html>.
- [104] Mitsubishi Motors. Mitsubishi asv-2 receives mlit type approval. Website, 2007. <http://media.mitsubishi-motors.com/pressrelease/e/corporate/detail458.html>.
- [105] El Mundo. Sistema de monitorización del nivel de alerta humano basado en ondas cerebrales. Prensa escrita, Octubre 1, 2007. Cuaderno 2: Tecnología.
- [106] Cristianini N. and Shawe-Taylor J. *An introduction to Support Vector Machines and other kernel-based learning methods*. Cambridge University Press, 2006.
- [107] Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans. Systems, Man and Cybernetics*, pages 62–66, 1979.
- [108] NHTSA. Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management. Final Report DOT HS 808 762, National Highway Traffic Safety Administration, Virginia 22161, USA, April 1998.
- [109] NHTSA. National highway traffic safety administration. Website, 2004. <http://www.nhtsa.dot.gov/>.

-
- [110] Federation of French motorway and toll facility companies (ASFA). Website, 2008. <http://www.autoroutes.fr/>.
- [111] International Organization of Motor Vehicle Manufacturers. Oica. Website, 2007. <http://www.oica.net/>.
- [112] Reporte Oriente. Los 10 países que producen más vehículos en el mundo. Website, 2008. <http://www.reporteoriente.wordpress.com/2008/03/25/los-10-paises-que-producen-mas-vehiculos-en-el-mundo/>.
- [113] Carrasco J. P., Escalera A. de la, and Armingol J. M. Speed supervision for intelligent vehicles. *Lecture Notes in Computer Science*, 4739:1175–1182, 2007.
- [114] Carrasco J. P., Escalera A. de la, and Armingol J. M. Speed supervision for intelligent vehicles. *11th. International Workshop on Computer Aided System Theory*, pages 399–402, February 12-16 2007. Las Palmas de Gran Canaria, Spain.
- [115] Gejgus P. and Sperka M. Face tracking in color video sequences. *Association for Computing Machinery*, pages 245–249, 2003.
- [116] Isasi P., Martínez P., and Borrajo D. *Lenguajes, gramáticas y autómatas: un enfoque práctico*. Lingüística computacional. Addison-Wesley, 2001.
- [117] Smith P., Shah M., and da Vitoria Lobo N. Determining driver visual attention with one camera. *IEEE Transactions on Intelligent Transportation Systems*, 4(4), 2003.
- [118] Viola P. and Jones M. Rapid object detection using a boosted cascade of simple features. *Computer Vision and Pattern Recognition, Proceedings of the 2001 IEEE Computer Society Conference on*, 1:1–511–1–518, 2001.
- [119] Viola P. and Jones M. Fast and robust classification using asymmetric adaboost and a detector cascade. *Advances in Neural Information Processing System, MIT Press, Cambridge, M. A.*, (14), 2002.
- [120] Viola P. and Jones M. Robust real-time object detection. *International Journal of Computer Vision - to appear*, 2002.

BIBLIOGRAFÍA

- [121] El Pais. Los ojos rojos sirven para contar miradas. Prensa escrita, Febrero 6, 2008. Tecnología: Sensores.
- [122] Immortal Project. Driver health and crash involvement. Website, 2005. <http://www.immortal.or.at>.
- [123] Ji Q. and Yang X. Real-time visual cues extraction for monitoring driver vigilance. *Lectures Notes in Computer Science, Proceedings of the Second International Workshop on Computer Vision Systems*, 2095:107–124, 2001.
- [124] Ji Q. and Yang X. Real-time eye, gaze and face pose tracking for monitoring driver vigilance. *Elsevier Science Ltd., Real Time Imaging*, 1(8):357–377, 2002.
- [125] Ji Q., Zhu Z., and Lan P. Real time nonintrusive monitoring and prediction of driver fatigue. *IEEE, Transaction on Vehicular Technology*, 53(4), Junio 2004.
- [126] Wang Q. and Yang J. Eye detection in facial images with unconstrained background. *Journal of Pattern Recognition Reseach*, 1:55–62, 2006.
- [127] Arellano-Valle R., Gómez M., and Quintana F. A new class of skew normal distributions. *PUC Techninical Report*, 2003.
- [128] Durrett R. *Probability: Theory and Examples*, volume 1. Library of Congress Catalogingin-Publication Data, 1991.
- [129] Grace R., Byrne V. E., Bierman D. M., Legrand J., Grcourt D., Davis R., Staszewski J., and Carnahan B. A drowsy driver detection system for heavy vehicles. *IEEE, Proceedings of Digital Avionics System Conference*, 2:1–8, October 1998.
- [130] Knipling R. and Wierwille W. Vehicle-based drowsy driver detection: Current status and future prospects. *IVSH America Fourth Annual Meeting*, 1994.
- [131] Parker J. R. *Practical Computer Vision Using C*. Wiley Professional Computing. John Wiley & Sons, Inc., 1994.
- [132] RACE. Datos generales de siniestralidad. Website, 2007. <http://www.race.es/portal/transform.jsp?seccion=/docs/20060110/0020.xml&xml=/docs/20060313/0001.xml&xsl=/contenido.xsl&menu=0&submenu=0&menu3=1>.

-
- [133] Grewal M. S. and Andrews A. P. *Kalman Filtering, Theory and Practice Using Matlab*. John Wiley & Sons Inc., 2001.
- [134] Hamlaoui S. and Davoine F. Facial action tracking using an AAM-based condensation approach. *IEEE, ICASSP05*, (2):701–704, June 2005.
- [135] Huang R. S., Kuo C. J., Tsai L. L., and Chen O. T. C. EEG pattern recognition-arousal states detection and classification. *IEEE Proceedings of, International Conference on Neural Networks*, 2:641–646, June 1996.
- [136] Milborrow S. and Nicolls F. Locating facial features with an extended active shape model, 2008. <http://www.milbo.users.sonic.net/stasm>.
- [137] National Semiconductor. *LM1881 - Video Sync Separator*, 2008. <http://www.national.com/mpf/LM/LM1881.html>.
- [138] Bando T., Shibata T., Doya K., and Ishii S. Switching particle filters for efficient visual tracking. *Robotics and Autonomous Systems*, 1(54):873–884, 2006.
- [139] Brandt T., Stemmer R., Mertsching B., and Rakotonirainy A. Affordable visual driver monitoring system for fatigue and monotony. *IEEE International Conference on Systems, Man and Cybernetics*, 7:6451–6456, 2004.
- [140] D’Orazio T., Leo M., and Distanto A. Eye detection in face images for a driver vigilance system. *IEEE, Intelligent Vehicle Symposium*, pages 95–98, June 2004.
- [141] Evgeniou T., Pontil M., Papageorgiou C., and Poggio T. Image representations for object detection using kernel classifiers. *In Asian Conference on Computer Vision*, 2000.
- [142] Hayami T., Matsunaga K., Shidoji K., and Matsuki Y. Detecting drowsiness while driving by measuring eye movement - a pilot study. *IEEE International Conference on Intelligent Transportation Systems*, pages 156–161, September 2002.
- [143] Lin C. T., Liang S. F., Chen Y. C., Hsu Y. C., and Ko L-W. Driver’s drowsiness estimation by combining eeg signal analysis and ica-based fuzzy neural networks. *IEEE International Symposium on Circuits and Systems, ISCAS*, May 2006.

BIBLIOGRAFÍA

- [144] Nabney I. T. *NetLab, Algorithms for Pattern Recognition*. Springer-Verlang, 2004 edition = 4ta.
- [145] AssistWare Technology. Safetrac. Website, 2007. <http://www.assistware.com/>.
- [146] Deutsche Verkehrswacht (DVW). Medien & service-center gmbh. Website, 2008. <http://www.dvw-ev.de/index.php>.
- [147] Ertico ITS Europe. Electronic warning system to keep drivers alert. Website, 2008. http://www.ertico.com/en/news_and_events/ertico_newsroom/electronic_warning_system_to_keep_drivers_alert.htm.
- [148] Dirección Nacional de Tráfico. Investigación de accidentes de peatones en carretera. Website, 2004. <http://www.dgt.es/educacionvial/recursos/dgt/EduVial/50/72/index.htm>.
- [149] Medicina tv.com. El sueño provoca alrededor de 40.000 accidentes de tráfico al año en españa. Website, 2007. <http://salud.medicinatv.com/noticias/default.asp?codigo=244296>.
- [150] Univision. Enfermedades mortales del futuro. Website, 2008. <http://www.univision.com/content/content.jhtml?chid=2&cid=1543723&schid=8241&secid=8261>.
- [151] Chen Y. W. and Kubo K. A robust eye detection and tracking technique using gabor filters. *IEEE, Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 1:109–112, November 2007.
- [152] Dong W. and Wu X. Driver fatigue detection based on distant eyelid. *IEEE, Int. Workshop VLSI Design & Video Tech.*, 2005.
- [153] Horng W., Chen C., and Chang Y. Driver fatigue detection based on eye tracking and dynamic template matching. *IEEE Proceedings of, International Conference on Networking, Sensing and Control*, 2004.

-
- [154] Rongben W., Keyou G., Shuming S., and Jiangwei C. A monitoring method of driver fatigue behavior based on machine vision. *IEEE, Proceedings on Intelligent Vehicles Symposium*, pages 110–113, June 2003.
- [155] Liu X., Xu F., and Fujimura K. Real-time eye detection and tracking for driver observation under various light conditions. *IEEE Intelligent Vehicle Symposim*, 2:344–351, June 2002.
- [156] Bar-Shalom Y. and Fortmann T. *Tracking and Data Association*, volume 119 of *Mathematics in Science and Engineering*. Academic Press Limited, Inc., 1988.
- [157] Freund Y. and Schapire R. A decision-theoretical generalization of online learning and an application to boosting. *In Second European Conference on Computational Learning Theory*, 1995.
- [158] Lee Y. Face detection and recognition using linear classification. Master’s thesis, Calgary University, 2004.
- [159] Liang Y., Reyes M. L., and Lee J. D. Real-time detection of driver cognitive distraction using support vector machines. *IEEE Transactions on Intelligent Transportation Systems*, 8(2):340–351, June 2007.
- [160] Tian Y., Kanade T., and Cohn J. F. Dual-state parametric eye tracking. *In Submit to International Conference on Face and Gesture Recognition*, 1999.
- [161] Wu Y., Liu H., and Zha H. A new method of detection humand eyelids based on deformable templates. *IEEE International Conference on Systems, Man and Cybernectics*, pages 604–609, June 2004.
- [162] Li S. Z., Chu R., Liao S., and Zhang L. Illumination invariant face recognition using near-infrared images. *IEEE Transactions on Pattern Recognition Analysis and Machine Intelligence*, 29(4):627–639, April 2007.
- [163] Tian Z. and Qin H. Real-time driver’s eye state detection. *IEEE, International Conference on Vehicular Electronics and Safety*, pages 285–289, October 2005.

BIBLIOGRAFÍA

- [164] Zhu Z., Fujimura K., and Ji Q. Real-time eye detection and tracking under various light conditions. *Proceedings of the 2002 Symposium of Eye tracking research & applications*, pages 139–144, 2002.
- [165] Zhu Z., Ji Q., Fujimura K., and Lee K. Combining Kalman filtering and mean shift for real time eye tracking under active ir illumination. *Proceedings of the 16 th International Conference on Pattern Recognition (ICPR'02)*, 4:318–321, 2002.